

The Naming Game on the complete graph

Eric Foxall

Abstract

We consider a model of language development, known as the naming game, in which agents invent, share and then select descriptive words for a single object, in such a way as to promote local consensus. When formulated on a finite and connected graph, a global consensus eventually emerges in which all agents use a common unique word. Previous numerical studies of the model on the complete graph with n agents suggest that when no words initially exist, the time to consensus is of order $n^{1/2}$, assuming each agent speaks at a constant rate. We show rigorously that the time to consensus is at least $n^{1/2-o(1)}$, and that it is at most constant times $\log n$ when only two words remain. In order to do so we develop sample path estimates for quasi-left continuous semimartingales with bounded jumps.

1 Introduction

The study of social dynamics from the standpoint of statistical physics is an area which has seen increased attention in recent years [4]. Historically, interacting particle system models of opinion dynamics, such as the voter model, have been of interest to mathematicians and studied in detail. However, new models emerging in the physics literature have yet to be given a fully rigorous mathematical treatment. One of these is a model of language development known as the naming game. This is a simple model of invention, sharing, and selection of words that displays eventual consensus towards a common vocabulary. It has been studied, using numerical simulations and heuristic computations, on lattices [1], the complete graph [3] and some random graphs [5]. As a first effort from the standpoint of probability theory, we study the naming game on the complete graph and give rigorous proof of some scaling relations that have been observed numerically.

We first recall the definition of the naming game on a general locally finite undirected graph $G = (V, E)$. Individuals correspond to vertices of the graph, and each individual speaks to its neighbours at a certain rate. The idea is that individuals are attempting to agree on a word to describe a certain object, for which initially, no descriptive words exist. The interaction rules are as follows.

- Speaker:
 - If the speaker does not know a word to describe the object then she invents a word and speaks it to the listener.
 - On the other hand, if the speaker does know at least one word to describe the object then she selects a word uniformly at random from her vocabulary and speaks it to the listener.
- Listener:
 - If the listener already knows the chosen word, then both speaker and listener delete the remainder of their vocabulary and remember only that word.
 - Otherwise, the listener adds the chosen word to their vocabulary.

Thus there is a mechanism both for the creation of new words, and for deletion and eventual agreement upon a single word. We now make this description rigorous. The process is denoted $(W_t)_{t \geq 0}$ with $W_t : V \rightarrow \mathcal{P}_o(V)$ for each $t \geq 0$, where $\mathcal{P}_o(V)$ is the collection of finite subsets of V . Thus, for each vertex $v \in V$, we have a process $W_t(v)$ whose state space consists of all finite subsets of the vertex set V and which is defined as

$$W_t(v) = \{w \in V : v \text{ knows the word invented by } w\}.$$

The process evolves as follows: For each $v \in V$, at the times of an independent Poisson process with rate one, v chooses a listener w uniformly at random from the set $\{u : uv \in E\}$; say this occurs at time t .

- If $W_{t-}(v)$ is empty then v speaks word v to w , so that $W_t(v) = \{v\}$ and $W_t(w) = W_{t-}(w) \cup \{v\}$.
- If $W_{t-}(v)$ is non-empty then v chooses a uniform random word u from $W_{t-}(v)$ and speaks it to w .
 - If $u \in W_{t-}(w)$ then $W_t(v) = W_t(w) = \{u\}$.
 - If $u \notin W_{t-}(w)$ then $W_t(v)$ is unchanged and $W_t(w) = W_{t-}(w) \cup \{u\}$.

If G is connected and finite, then with probability one, the system eventually settles into one of the set of absorbing states

$$\{W_t(v) = \{w\} \text{ for all } v \in V : w \in V\}$$

and we would like to know what happens on the way to this consensus. Let

$$V_t = \bigcup_v W_t(v)$$

denote the set of words in existence at time t . If G is the complete graph on n vertices, i.e.,

$$V = \{1, \dots, n\} \quad \text{and} \quad E = \{\{v, w\} : v, w \in V, v \neq w\},$$

numerical studies and heuristic computations [2] indicate three distinct phases.

1. Early phase: V_t rises from 0 to about $n/2$ in about $\frac{1}{2} \log n$ time.
2. Middle phase: V_t remains fairly constant up till about $n^{1/2}$ time.
3. Late phase: V_t falls sharply to 1 within about $n^{1/4}$ time.

In this article we consider the early and middle phases, and what we call the final phase, where we assume that V_t is initially equal to 2, and track the dynamics until it goes to 1. The bulk of the late phase, during which the diversity of language collapses from a large number to a small number of different words, is more difficult to assess, and is not considered here.

In the next section we construct the model as a stochastic process, then describe the main results and give the layout for the rest of the article.

2 Construction and Main Results

We first note a useful “graphical construction” of the process, on a general locally finite graph G , from arbitrary initial data. We assume the vertices are totally ordered according to some fixed order. Given $\mu > 0$, let $\{(s_i, u_i) : i \geq 1\}$ be an independent and identically distributed sequence, with each s_i exponentially distributed with mean one and each u_i independent of s_i and uniform on $[0, 1]$, and for $i \geq 1$, let $t_i = \mu^{-1} \sum_{j=1}^i s_j$. Then, the set of points

$$U := \{(t_i, u_i) : i \geq 1\} \subset \mathbb{R}_+ \times [0, 1]$$

defines what we call an augmented Poisson point process with intensity μ , since (t_i) are the jump times of a Poisson process with intensity μ and each point t_i comes equipped with an independent uniform random variable u_i to help with the decision-making process.

Let F denote the set of directed edges $\{(v, w) : vw \in E\}$, and associate to each directed edge $(v, w) \in F$ an independent augmented Poisson point process $U(v, w)$ with intensity $(\deg v)^{-1}$. Suppose that $(t, u) \in U(v, w)$ and $|W_{t-}(v)| = k$, with $W_{t-}(v) = \{w_1, \dots, w_k\}$ labelled in increasing order.

- If $k = 0$ then v speaks word v to w at time t .
- If $k \geq 1$, then v speaks word w_i to w at time t if and only if

$$(i-1)/k \leq u < i/k.$$

We then follow the rules as described above to determine W_t . If G is a finite graph, then since the intensity of the union $\bigcup_{(v,w) \in F} U(v,w)$ is finite, its points are well-ordered in time with probability 1, and so W_t can be determined from the initial state and the points $U(v,w)$ by updating sequentially in time. If G is an infinite graph, one needs to ensure that for each spacetime point (v,t) , a finite number of events suffices to determine $W_t(v)$. Although this is not hard to do, we will ignore it since from here on we focus on the case where G is the complete graph on n vertices and thus finite for any n .

Recall that $V_t = \bigcup_v W_t(v)$ denotes the set of words in existence at time t . The following result gives estimates of V_t in the middle phase of the process.

Theorem 1. *For any $\epsilon > 0$, let $a = (\frac{1}{2} + \epsilon) \log n$ and $b = n^{1/2-\epsilon}$. Then as $n \rightarrow \infty$*

$$\mathbb{P}(\sup_{a \leq t \leq b} |V_t - \frac{n}{2}| = o(n)) = 1 - o(1).$$

The result is proved in two main steps.

1. First, we show that $n/2 + n^{1/2+o(1)}$ are ever created, and within $(\frac{1}{2} + o(1)) \log n$ time.
2. Then, we show that $o(n)$ words are deleted in $n^{1/2-o(1)}$ time.

The proof relies on approximating the size of the *cluster* $\mathcal{C}_t(w)$ corresponding to a given word w by a sort of branching process evolving in a non-stationary random environment. The cluster is defined by

$$\mathcal{C}_t(w) = \{v : w \in W_t(v)\}$$

and is the set of individuals that know word w at time t . We also need to control the correlation between distinct clusters $\mathcal{C}_t(w_1), \mathcal{C}_t(w_2)$. To achieve both tasks we will use a slightly modified graphical construction which is better tailored to tracking the evolution of one or more distinguished clusters.

For the next result we introduce some notation. Let $\Theta_t(W)$ denote the configuration at time t when the initial configuration is W , and let $V' \subset V$. If $W : V \rightarrow \mathcal{P}_o(V') \setminus \{\emptyset\}$ then clearly

$$\Theta_t(W) : V \rightarrow \mathcal{P}_o(V') \setminus \{\emptyset\} \quad \text{for all } t > 0,$$

that is, if each vertex has initially a non-empty vocabulary consisting of words in V' , the same is true at later times. In particular, if $V' = \{A, B\}$ for a pair of words A, B , then each vertex has one of the three types A, B and AB . We note that, starting from $W(v) = \emptyset$ for all v , before the process achieves consensus there is a good chance that at some point only two words remain, so we can think of it as the final phase of the process. For the complete graph on n vertices, the rate of change of the number of individuals of each type does not depend on the particular location of the individuals. Therefore, letting X_t, Y_t, Z_t denote the number of sites at time t with respective types A, B and AB , the process $\Phi_t = (X_t, Y_t, Z_t)$ is a continuous-time Markov chain. Since the states $\Phi^A = (n, 0, 0)$ and $\Phi^B = (0, n, 0)$ are the only absorbing states and are both accessible from all other states, it follows that with probability one,

$$\lim_{t \rightarrow \infty} \Phi_t \in \{\Phi^A, \Phi^B\}.$$

The following result characterizes how long this takes, for large n . Use $\mathbb{P}_{(X,Y,Z)}(\cdot)$ for the law of the process with initial configuration (X, Y, Z) .

Theorem 2. *Let $\gamma = 1 + (-4 + 2\sqrt{5})^{-1}$, and define the stopping time*

$$T_c = \inf \{t : \Phi_t \in \{\Phi^A, \Phi^B\}\}.$$

Then, for any $\alpha > 0$,

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_{(X,Y,Z)} \mathbb{P}_{(X,Y,Z)}(T_c / \log n > \gamma + \alpha) &= 0 \\ \lim_{n \rightarrow \infty} \sup_{(X,Y,Z)} \mathbb{P}_{(X,Y,Z)}(T_c / \log n > \gamma - \alpha) &= 1 \end{aligned}$$

Notice that, if individuals only remember the last word they heard, then starting from a configuration with two words, we obtain the voter model on the complete graph, for which the time to consensus is of order n . The reason it is much faster here is because, once a majority of type A or B develops, it is maintained. To prove this result we use an ODE heuristic to get an idea of what is happening, then carve up the state space into a few pieces and use martingale estimates to control the behavior of sample paths on each piece.

The paper is laid out as follows. In Section 3 we derive a simple and useful sample path estimate for quasi-left continuous semimartingales with bounded jumps, and give some formulas that help with computations later on. This section can be read independently of the rest of the paper, and may be of use in other applications. In Section 4 we prove Theorem 1 in several steps. In Section 4.1 we show that about $n/2$ words are created in about $\frac{1}{2} \log n$ time, using Chebyshev's inequality and a coupon-collecting argument, respectively. In Section 4.2 we show that $o(n)$ words are deleted in $n^{1/2-\epsilon}$ time, which as noted above is achieved by controlling the number of individuals that know a given word, and which requires the sample path estimates of Section 3. In Section 5 we use an ODE comparison and the estimates of Section 3 to prove Theorem 2. Some additional results are collected in an Appendix, including a general sample path estimate for Poisson processes, and one for semimartingales with sublinear drift.

3 Sample path estimation

Using the semimartingale theory in [6] we derive a useful estimate for quasi-left continuous semimartingales with bounded jumps, which can be found in Lemma 3. It can be thought of as a continuous-time analogue of Azuma's inequality. In this section, unless otherwise noted, references are to formulas in [6].

Given is a filtered probability space $(\Omega, \mathcal{F}, \mathbf{F}, P)$ satisfying the “usual conditions” as described in [6]. Processes are assumed to be optional. Given X , X_- is the left continuous process $(X_{t-})_{t \geq 0}$ and $\Delta X = X - X_-$ is the process of jumps. X^p denotes the compensator and X^c the continuous martingale part, when they exist.

A *semimartingale* X is a process (on \mathbb{R} unless specified otherwise) that can be written as $X = X_0 + M + A$, where X_0 is an \mathcal{F}_0 -measurable random variable, M is a local martingale and A has locally finite variation. Using I.3.17, a semimartingale is *special* if it can be written as

$$X = X_0 + X^m + X^p \tag{1}$$

where X^p is the compensator of X and X^m is a uniquely defined local martingale satisfying $X_0^m = 0$. If X is a semimartingale with *bounded jumps*, that is, $|\Delta X| \leq c$ for some $c > 0$ then by I.4.24, X is a special semimartingale and $|\Delta X^m| \leq 2c$. If X is also *quasi-left continuous*, that is, $\Delta X_T = 0$ a.s. on $\{T < \infty\}$, for any predictable time T , then using I.2.35 in the proof of I.4.24, we obtain the slightly stronger estimate $|\Delta X^m| \leq c$.

Any (right-continuous) Markov chain with values in \mathbb{R} is a semimartingale, since it is right-continuous and has locally finite variation, and is also quasi-left continuous, effectively because the jump times of a Poisson process are totally inaccessible; if this explanation is insufficient use Proposition 22.20 in [?] and note that Markov chains are Feller processes. As shown in I.4.28, a deterministic function $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ is a semimartingale iff it is right-continuous with finite variation over each compact interval, and is quasi-left continuous iff it is continuous, since any fixed time is predictable.

We will occasionally assume X is defined only up to some predictable time ζ that may be finite; in this case, information about X can be recovered from the stopped processes X^{τ_n} defined by $X_t^{\tau_n} = X_{t \wedge \tau_n}$, where τ_n is an announcing sequence for ζ , i.e., an increasing sequence of stopping times with limit ζ .

If M is a local martingale satisfying $M_0 = 0$ and $|\Delta M| \leq c$ for some $c > 0$, by I.4.1, M is locally square-integrable, so by I.4.3, M^2 has a compensator, denoted $\langle M \rangle$ and called the predictable quadratic variation.

Relative to the decomposition (I.4.18) $M = M^c + M^d$ into continuous and discontinuous martingale parts,

$$\langle M \rangle = \langle M^c \rangle + \langle M^d \rangle,$$

and $\langle M^d \rangle$ is the compensator of $[M^d] = \sum_{s \leq t} (\Delta M_s)^2$, the quadratic variation of M^d .

Lemma 1. *Let M be a quasi-left continuous local martingale with $M_0 = 0$ and $|\Delta M| \leq c$ for some $c > 0$. Then,*

$$\exp(M - (e^c/2)\langle M \rangle)$$

is a local supermartingale with initial value 1.

Proof. Let V_t be a continuous predictable process with locally finite variation, satisfying $V_0 = 0$, and let $E = \exp(M - V)$. Applying Itô's formula I.4.57 using the function $x \mapsto e^x$,

$$E_t = 1 + (E_- \cdot (M - V))_t + \frac{1}{2}(E_- \cdot \langle M^c \rangle)_t + \sum_{s \leq t} E_{s-} (e^{\Delta M_s} - 1 - \Delta M_s). \quad (2)$$

Noting that $e^x - 1 - x \leq \frac{1}{2}e^c x^2$ when $|x| \leq c$, the last term is bounded by

$$\frac{e^c}{2} \sum_{s \leq t} E_{s-} (\Delta M_s)^2.$$

Using this bound and taking the compensator of both sides in (2),

$$\begin{aligned} E^p &\leq E_- \cdot (-V + \frac{1}{2}\langle M^c \rangle + \frac{1}{2}e^c \langle M^d \rangle) \\ &\leq E_- \cdot (-V + \frac{1}{2}e^c \langle M \rangle). \end{aligned}$$

The assumption of quasi-left continuity implies that $\langle M \rangle$ can be taken continuous (i.e. it has a continuous version; see I.4.3). Since it is the compensator of $[M, M]$, $\langle M \rangle$ is also predictable and has locally finite variation. Letting $V = (e^c/2)\langle M \rangle$, the same is true for V . With this choice of V , $E^p \leq 0$, which implies E is a local supermartingale. \square

Lemma 2. *Let X be a special semimartingale with locally square-integrable martingale part X^m . Then, X is quasi-left continuous iff X^p and $\langle X^m \rangle$ are continuous.*

Proof. If X is quasi-left continuous (qlc), then by I.2.35 its predictable projection ${}^pX = X_-$. Since the ${}^p(\cdot)$ operation is linear (which follows from uniqueness and property (ii) in I.2.28), and since ${}^pX_- = X_-$, it follows that

$${}^p(\Delta X) = {}^pX - {}^pX_- = 0.$$

Since X is special it has a compensator X^p , and by I.3.21, $\Delta(X^p) = {}^p(\Delta X)$. By the above, this is 0, i.e., X^p is continuous. Using (1), $\Delta X^m = \Delta X$ which implies that X^m is qlc, by definition of qlc. Since X^m is locally square-integrable, by I.4.3, $\langle X^m \rangle$ is continuous.

On the other hand, if $\langle X^m \rangle$ is continuous then by I.4.3, X^m is qlc. If in addition X^p is continuous then $\Delta X = \Delta X^m$ which implies X is qlc, by definition of qlc. \square

Lemma 3. *Let X be a quasi-left continuous semimartingale such that $|\Delta X| \leq c$ for some $c > 0$. Then for $\lambda, a > 0$ and $\bullet \in \pm$,*

$$P(\bullet(X_t - X_0 - X_t^p) \geq a + (\lambda e^{\lambda c}/2)\langle X^m \rangle_t) \leq e^{-\lambda a} \quad (3)$$

Proof. Notice that $X_t - X_0 - X_t^p = X_t^m$ and that for $\lambda > 0$ and $\bullet \in \pm$, $\langle \bullet \lambda X^m \rangle = \lambda^2 \langle X^m \rangle$. As noted just above (1), since X has bounded jumps it is special and by Lemma 2, X^m is qlc. Take $M = \bullet \lambda X^m$ in Lemma 1, which has $|\Delta M| \leq \lambda c$, and use Doob's inequality to find

$$P(\bullet \lambda X_t^m - (\lambda^2 e^{\lambda c}/2)\langle X^m \rangle_t \geq \lambda a) \leq e^{-\lambda a}.$$

\square

For practicality's sake we'll use a slightly cruder version of (3). Since $1/2 \leq \log 2$, if $\lambda c \leq 1/2$ then $e^{\lambda c} \leq 2$, so from (3) it follows that for $a > 0$ and $\bullet \in \pm$,

$$\text{if } 0 < \lambda c \leq 1/2 \text{ then } P(\bullet(X_t - X_0 - X_t^p) \geq a + \lambda \langle X^m \rangle_t) \leq e^{-\lambda a}. \quad (4)$$

Using Lemma 2 as inspiration, say that a special semimartingale X with locally square-integrable martingale part X^m is *quasi-absolutely continuous* (qac) if both X^p and $\langle X^m \rangle$ are absolutely continuous. In this case define the *drift* $\mu(X) = (\mu_t(X))_t$ and the *diffusivity* $\sigma^2(X) = (\sigma_t^2(X))_t$ for Lebesgue-a.e. t by

$$\mu_t(X) = \frac{d}{dt} X_t^p, \quad \sigma_t^2(X) = \frac{d}{dt} \langle X^m \rangle_t. \quad (5)$$

For deterministic processes, qac is equivalent to absolute continuity, since $\mu_t(f) = f(t)$, $\sigma_t^2(f) = 0$ and absolute continuity implies locally finite variation. For Markov chains X on \mathbb{R} with jump measure $\alpha(x, dy)$, if qac holds then μ and σ are given by functions

$$\mu(x) = \int_{\mathbb{R}} y \alpha(x, dy), \quad \sigma^2(x) = \int_{\mathbb{R}} y^2 \alpha(x, dy),$$

i.e., $\mu_t(X) = \mu(X_t)$ and $\sigma_t^2(X) = \sigma^2(X_t)$. Conversely, if $|\Delta X| \leq c$ and the total intensity $q(x) = \int_{\mathbb{R}} \alpha(x, dy)$ of the jump measure is bounded on compact subintervals of \mathbb{R} , then X is qac up to the first explosion time $\sup_{r>0} \inf\{t : |X_t| \geq r\}$, and $\sigma^2(x) \leq c^2 q(x)$.

Lemma 4. [Product rule] Suppose X_t, Y_t are qac semimartingales on a common filtered probability space. Then $\langle X^m, Y^m \rangle$ exists and is absolutely continuous, and $(XY)^p$ exists, is absolutely continuous, and $\mu_t(XY) = \frac{d}{dt} (XY)^p_t$ is given by

$$\mu(XY) = \sigma(X, Y) + X_- \mu(Y) + Y_- \mu(X),$$

where $\sigma_t(X, Y) = \frac{d}{dt} \langle X^m, Y^m \rangle_t$.

Proof. By definition of quadratic variation,

$$XY = X_0 Y_0 + [X, Y] + X_- \cdot Y + Y_- \cdot X.$$

Since $X = X_0 + X^p + X^m$, $Y = Y_0 + Y^p + Y^m$ and $X_0 + X^p$, $Y_0 + Y^p$ have locally finite variation, $[X, Y] = [X^m, Y^m]$. Since X^m, Y^m are locally square-int, $[X^m, Y^m]$ has compensator $\langle X^m, Y^m \rangle$, so XY has compensator

$$(XY)^p = \langle X^m, Y^m \rangle + X_- \cdot Y^p + Y_- \cdot X^p.$$

The result will follow if we can show $\langle X^m, Y^m \rangle$ is absolutely continuous. For any $s < t$, applying the Cauchy-Schwarz inequality to the symmetric, bilinear and semidefinite map $(X, y) \mapsto \langle X, Y \rangle_t - \langle X, Y \rangle_s$ gives

$$|\langle X^m, Y^m \rangle_t - \langle X^m, Y^m \rangle_s| \leq \sqrt{(\langle X^m \rangle_t - \langle X^m \rangle_s)(\langle Y^m \rangle_t - \langle Y^m \rangle_s)}.$$

Absolutely continuity of $t \mapsto \langle X^m \rangle_t, \langle Y^m \rangle_t$ means that for any $\epsilon > 0$ there is $\delta > 0$ so that if $\sum_i |t_i - s_i| < \delta$ then $\sum_i |\langle X^m \rangle_{t_i} - \langle X^m \rangle_{s_i}|, \sum_i |\langle Y^m \rangle_{t_i} - \langle Y^m \rangle_{s_i}| < \epsilon$. Using the Cauchy-Schwarz inequality to obtain the second line,

$$\begin{aligned} \sum_i |\langle X^m, Y^m \rangle_{t_i} - \langle X^m, Y^m \rangle_{s_i}| &\leq \sum_i \sqrt{(\langle X^m \rangle_{t_i} - \langle X^m \rangle_{s_i})(\langle Y^m \rangle_{t_i} - \langle Y^m \rangle_{s_i})} \\ &\leq \left(\sum_i |\langle X^m \rangle_{t_i} - \langle X^m \rangle_{s_i}| \sum_i |\langle Y^m \rangle_{t_i} - \langle Y^m \rangle_{s_i}| \right)^{1/2} \\ &< (\epsilon \cdot \epsilon)^{1/2} = \epsilon \end{aligned}$$

which shows that $\langle X^m, Y^m \rangle$ is absolutely continuous. \square

4 Early and middle phases

In this section we consider the behaviour of $|V_t|$ for $t \leq n^{1/2-o(1)}$. Define

$$V_t^o = \bigcup_{(v,s): s \leq t} W_s(v) \quad \text{and} \quad V_t^\times = V_t \setminus V_t^o,$$

respectively the number of words created up to time t , and the number of words created and then deleted by time t . Theorem 1 is implied by the following two propositions, whose proof is the objective of this section.

Proposition 1. *For each $\epsilon > 0$, $\lim_{n \rightarrow \infty} \mathbb{P}(\sup_{t \geq (\frac{1}{2} + \epsilon) \log n} |V_t^o| - \frac{n}{2} \geq n^{1/2+\epsilon}) = 0$.*

Proposition 2. *For each $\epsilon > 0$, $\lim_{n \rightarrow \infty} \mathbb{P}(\sup_{t \leq n^{1/2-\epsilon}} |V_t^\times| = o(n)) = 1$.*

In words, in order to estimate $|V_t|$ we obtain good control on $|V_t^o|$, then show that $|V_t^\times|$ is not too big. We begin with V_t^o .

4.1 Creation of vocabulary

Our first task is to prove Proposition 1, and to do so we show that $|V_t^o|$ rises from 0 to $n/2 + O(n^{1/2+o(1)})$ within $\frac{1}{2} \log n$ time, then remains constant. For a vertex v let $N_t(v) = |W_t(v)|$ denote the size of the vocabulary of individual v , and let

$$T_o = \inf\{t : \min_v N_t(v) \geq 1\}$$

be the first time that every individual knows at least one word. Clearly V_t^o is non-decreasing as a set, so $V_\infty^o = \lim_{t \rightarrow \infty} V_t^o$ exists and $|V_\infty^o| \leq n$. Once everyone knows a word, no new words are created, so $V_t = V_{T_o}^o = V_\infty^o$ for $t \geq T_o$. Proposition 1 is implied by the following two lemmas, in which we estimate T_o and $V_{T_o}^o$.

Lemma 5. *For $c \geq 0$,*

$$\mathbb{P}(|T_o - \frac{1}{2} \log n| \geq c) \leq 2e^{-c} + o(1) \text{ as } n \rightarrow \infty.$$

Lemma 6. *Let $X = |V_{T_o}^o|$ be the number of words ever created. Then,*

$$\lim_{n \rightarrow \infty} P(|X - n/2| \geq n^\alpha) = 0 \text{ for all } \alpha > 1/2.$$

Proof of Lemma 5. Let $M_t = \{v : N_t(v) = 0\}$ denote *mute* vertices, those not yet knowing a word, and observe that $T_o \leq t$ is equivalent to $|M_t| = 0$. For each distinct ordered pair of vertices (v, w) , at rate $(n-1)^{-1}$, the directed edge (v, w) has an event, and both v and w are removed from M_t , if either or both still belongs. If we let $Z_t = |M_t|$ denote the number of mute vertices at time t , it follows that Z_t is a Markov chain with $Z_0 = n$ and transitions

$$Z_t \rightarrow \begin{cases} Z_t - 1 & \text{at rate } 2(n-1)^{-1}Z_t(n-Z_t), \text{ and} \\ Z_t - 2 & \text{at rate } (n-1)^{-1}Z_t(Z_t-1). \end{cases}$$

We find that

$$\begin{aligned} \lim_{h \rightarrow 0^+} h^{-1} \mathbb{E}[Z_{t+h} - Z_t \mid Z_t = z] &= -2(n-1)^{-1}z(n-z) - 2(n-1)^{-1}z(z-1) \\ &= -2(n-1)^{-1}(nz - z^2 + z^2 - z) \\ &= -2(n-1)^{-1}(n-1)z = -2z. \end{aligned}$$

Letting $m(t) = \mathbb{E}[Z_t]$, $m(0) = n$ and taking expectations in the above, $m'(t) = -2m(t)$, which has the unique solution $m(t) = ne^{-2t}$. Fix $c \in \mathbb{R}$ and let $t_c = \frac{1}{2} \log n + c$. Using Markov's inequality,

$$\mathbb{P}(T_o > t_c) = \mathbb{P}(Z_{t_c} \geq 1) \leq \mathbb{E}[Z_{t_c}] = e^{-2c}.$$

To get a lower bound we turn to Z_t^2 , which has transitions

$$Z_t^2 \rightarrow \begin{cases} Z_t^2 - Z_t + 1 & \text{at rate } 2(n-1)^{-1}Z_t(n-Z_t), \text{ and} \\ Z_t^2 - 4Z_t + 4 & \text{at rate } (n-1)^{-1}Z_t(Z_t-1), \end{cases}$$

so

$$\begin{aligned} \lim_{h \rightarrow 0^+} h^{-1} \mathbb{E}[Z_{t+h}^2 - Z_t^2 \mid Z_t = z] &= -(2z-1)2(n-1)^{-1}z(n-z) - (4z-4)(n-1)^{-1}z(z-1) \\ &= -4z(n-1)^{-1}((z-\frac{1}{2})(n-z) + (z-1)^2) \\ &= -4z(n-1)^{-1}(nz - z^2 - \frac{n}{2} + \frac{z}{2} + z^2 - 2z + 1) \\ &= -4z(n-1)^{-1}((n-\frac{3}{2})z + 1 - \frac{n}{2}) \\ &= \frac{2(n-2)}{n-1}z - \frac{4(n-3/2)}{n-1}z^2. \end{aligned}$$

Letting $\nu(t) = \mathbb{E}[Z_t^2]$, $\nu(0) = n^2$ and taking expectations above,

$$\nu'(t) = -4(1 - (2(n-1))^{-1})\nu(t)^2 + 2(1 - (n-1)^{-1})m(t),$$

so letting $\gamma = 4 - 2/(n-1)$, using $m(t) = ne^{-2t}$ and solving the above DE, we find

$$\nu(t) = n^2 e^{-\gamma t} + 2(1 - 1/(n-1))ne^{-\gamma t}(e^{(\gamma-2)t} - 1)/(\gamma - 2).$$

As above let $t_c = \frac{1}{2} \log n + c$, then $m(t_c) = e^{-2c}$ and for fixed c ,

$$\nu(t_c) = e^{-4c} + e^{-2c} + o(1) \text{ as } n \rightarrow \infty,$$

so $\text{Var}(Z_{t_c}) = \nu(t_c) - m(t_c)^2 = e^{-2c} + o(1)$. Using Chebyshev's inequality,

$$\mathbb{P}(T_o \leq t_c) = \mathbb{P}(Z_t = 0) \leq \mathbb{P}(|Z_t - \mathbb{E}[Z_t]| \geq \mathbb{E}[Z_t]) \leq \text{Var}(Z_{t_c}) / \mathbb{E}[Z_{t_c}]^2 \leq \frac{e^{-2c} + o(1)}{e^{-4c}} = e^{2c} + o(1).$$

The result follows by taking a union bound of both estimates. \square

We note in passing that $|V_0^o| = 0$ and $|V_t^o|$ increases by 1 at rate Z_t . Heuristically, $Z_t \approx ne^{-2t}$, so $|V_t^o| \approx (n/2)(1 - e^{-2t})$, for $t \leq \frac{1}{2} \log n$. This can be made precise using stochastic calculus, although we do not pursue it here.

Proof of Lemma 6. Letting X_v for each vertex $v \in V$ be the Bernoulli random variable equal to one if and only if v speaks before listening, by construction and obvious symmetry, we have

$$X = \sum_{v \in V} X_v \quad \text{and} \quad P(X_v = 0) = P(X_v = 1) = 1/2.$$

It follows that the expected number of words is given by

$$E(X) = \sum_{v \in V} E(X_v) = \sum_{v \in V} P(X_v = 1) = n/2. \quad (6)$$

To also compute the variance, fix $v, w \in V$ and let B be the event that the first edge becoming active starting from v or w is edge vw . Since there are $n-1$ edges starting from each vertex,

$$P(B) = \frac{1}{2(n-1)-1} = \frac{1}{2n-3}. \quad (7)$$

In addition, the two vertices cannot both speak before listening when B occurs whereas the two events are independent on the event B^c therefore

$$\begin{aligned} P(X_v = X_w = 1 \mid B) &= 0 \\ P(X_v = X_w = 1 \mid B^c) &= P(X_v = 1 \mid B^c) P(X_w = 1 \mid B^c) = 1/4. \end{aligned} \quad (8)$$

Combining (7)–(8), we deduce that

$$\begin{aligned} E(X^2) &= \sum_{v \in V} P(X_v^2 = 1) + \sum_{v \neq w} P(X_v = X_w = 1) \\ &= \sum_{v \in V} \frac{1}{2} + \sum_{v \neq w} \frac{1}{4} \frac{2n-4}{2n-3} = \frac{n}{2} \left(1 + \frac{(n-1)(n-2)}{2n-3} \right) \end{aligned}$$

which, together with some basic algebra, gives the variance

$$\text{Var}(X) = \frac{n}{2} \left(1 + \frac{(n-1)(n-2)}{2n-3} - \frac{n}{2} \right) = \frac{n}{4} \left(\frac{n-2}{2n-3} \right) = O(n). \quad (9)$$

From (6) and (9) and Chebyshev's inequality, we conclude that

$$\lim_{n \rightarrow \infty} P(|X - n/2| \geq n^\alpha) \leq \lim_{n \rightarrow \infty} n^{-2\alpha} \text{Var}(X) = 0$$

for all $\alpha > 1/2$. This completes the proof. \square

4.2 Maintenance of vocabulary

Next, we prove Proposition 2, that says that with probability tending to 1 as $n \rightarrow \infty$,

$$\sup_{t \leq n^{1/2-o(1)}} |V_t^\times| = o(n).$$

Clearly V_t^\times , like V_t^o , is non-decreasing, since once a word vanishes from the population, it does not come back. We first bound $|V_t^\times|$ by a simpler quantity. Say that *agreement* upon word y occurs at (v, w, t) if

$$y \in W_{t-}(w) \text{ and } v \text{ speaks word } y \text{ to } w \text{ at time } t.$$

If word w is created at some time $s \leq t$, then $w \in W_s(w)$, and remains in individual w 's vocabulary at least until the first time $t > s$ that agreement occurs at (\cdot, w, t) or (w, \cdot, t) . This implies

$$V_t^\times \subseteq H_t = \{w : \text{agreement occurs at } (\cdot, w, s) \text{ or } (w, \cdot, s) \text{ for some } s \leq t\}.$$

In words, in order to delete a word w from the population, it must at least be deleted from its source. Since each agreement contributes at most 2 to H_t , it follows that

$$|V_t^\times| \leq 2A_t \quad \text{where} \quad A_t = \frac{|\{s \leq t : \text{agreement occurs at } (\cdot, \cdot, s)\}|}{(\text{number of agreements up to time } t)}.$$

In order to control A_t we first define some useful observable quantities. For $w \in V$ we recall the *cluster* $\mathcal{C}_t(w)$ of w , that is, the set of individuals that know word w at time t :

$$\mathcal{C}_t(w) = \{v : w \in W_t(v)\}.$$

Recall that $N_t(v) = |W_t(v)|$ denotes the size of the vocabulary of individual v , and let

$$R_t(w) = \mathbf{1}(N_t(w) = 0) + \sum_{v \in \mathcal{C}_t(w)} 1/N_t(v) \quad (10)$$

denote the rate at which word w is spoken. Let $J(w, v)$ denote the times at which w speaks to v , and let

$$N_t^\ell(v) = \sum_w |J(w, v) \cap [0, t]| = \text{number of listening events for } v \text{ up to time } t,$$

noting that $N_t(v) \leq N_t^\ell(v)$ and $\{(N_t^\ell(v)) : v \in V\}$ is a collection of independent Poisson processes with intensity 1. If we let

$$\begin{aligned} \tau_a(v) &= \inf\{t : v \in H_t\} \quad \text{and} \\ \tau_a(v, t) &= 0 \vee \sup\{s \leq t : \text{agreement occurs at } (v, \cdot, s) \text{ or } (\cdot, v, s)\}, \end{aligned}$$

then $N_t(v) = N_t^\ell(v) - N_{\tau_a(v,t)-}^\ell(v)$, and in particular,

$$N_t(v) = N_t^\ell(v) \text{ for } t < \tau_a(v).$$

Let $S_t(w) = |\mathcal{C}_t(w)|$ and $P_t(w) = (S_t(w) - 1)/(n - 1)$, and let $S_t = \max_w S_t(w)$. Each site v that knows word w speaks it at rate $N_t(v)^{-1}/(n - 1)$ to each of the other $S_t(w) - 1$ sites in $\mathcal{C}_t(w)$. Letting

$$A_t(w) = |\{s \leq t : \text{agreement occurs upon word } w \text{ at time } s\}|,$$

so that $A_t = \sum_w A_t(w)$, it follows that $A_t(w)$ increases by 1 at rate

$$(S_t(w) - 1) \sum_{v \in \mathcal{C}_t(w)} \frac{N_t(v)^{-1}}{n - 1} = R_t(w)P_t(w).$$

Since $\sum_{w \in V} R_t(w) = n$ is the total speaking rate and $P_t(w) \leq (S_t - 1)/(n - 1) \leq S_t/n$, summing the above display over $w \in V$ we find

$$A_t \text{ increases by 1 at rate at most } S_t. \quad (11)$$

We have reduced the problem of controlling $|V_t^\times|$ to that of controlling S_t . The following becomes the goal of this subsection. Since its proof has a few parts, we call it a theorem.

Theorem 3. *For small $\epsilon > 0$,*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\sup_{t \leq n^{1/2-\epsilon}} \frac{S_t}{(1+t)^{1+\epsilon}} \geq (\log n)^9\right) = 0.$$

Before moving onto the proof of Theorem 3 we first use it to obtain Proposition 2.

Proof of Proposition 2. From (11), for any $T > 0$, $\sup_{t \leq T} A_t \leq \text{Poisson}(\int_0^T S_u du)$. Using Theorem 3, with probability $1 - o(1)$ as $n \rightarrow \infty$

$$\begin{aligned} \int_0^{n^{1/2-\epsilon}} S_u du &\leq (\log n)^9 \int_0^{n^{1/2-\epsilon}} (1+u)^{1+\epsilon} du \leq (\log n)^9 (2+\epsilon)^{-1} (1+n^{1/2-\epsilon})^{2+\epsilon} \\ &= O((\log n)^9 n^{1-3\epsilon/2-\epsilon^2}) = o(n). \end{aligned}$$

Since $\mathbb{P}(\text{Poisson}(\lambda) \leq 2\lambda) \rightarrow 1$ as $\lambda \rightarrow \infty$ it follows that $\sup_{t \leq n^{1/2-\epsilon}} A_t = o(n)$ with probability $1 - o(1)$, and since $|V_t^\times| \leq 2A_t$, the same is true for $|V_t^\times|$. \square

To begin the proof of Theorem 3 we introduce a modified construction to help us make a coupling. First, for each ordered triple (y, z, v) let $R_t(y, z, v)$ be the rate at which word y is spoken by site z to v , let $R_t(y, v) = \sum_z R_t(y, z, v)$ be the rate at which site v hears word y , and as above let $R_t(y) = \sum_v R_t(y, v)$ be the rate at which word y is spoken. We calculate

$$\begin{aligned} R_t(y, z, v) &= (N_t(z))^{-1} \mathbf{1}(z \in \mathcal{C}_t(y), z \neq v) + \mathbf{1}(y = z \neq v, N_t(y) = 0)/(n - 1) \text{ and} \\ R_t(y, v) &= (\mathbf{1}(N_t(y) = 0, v \neq y) + \sum_{z \in \mathcal{C}_t(y) \setminus \{v\}} N_t(z)^{-1})/(n - 1). \end{aligned} \quad (12)$$

Clearly $\sum_y R_t(y, v) = 1$ for each v, w and $t \geq 0$. Fix an ordering $v_1 < \dots < v_n$ of V and define an independent family $\{U_v : v \in V\}$ of augmented Poisson point processes with intensity 1, that will correspond to listening events. For $v \in V$, $1 \leq i, j \leq n$ and $t \geq 0$ let

$$I_t(v, i, j) = \left[\sum_{k=1}^{i-1} R_t(v_k, v) + \sum_{m=1}^{j-1} R_t(v_i, v_m, v), \sum_{k=1}^{i-1} R_t(v_k, v) + \sum_{m=1}^j R_t(v_i, v_m, v) \right),$$

noting that $\{I_t(v, i, j) : 1 \leq i, j \leq n\}$ partitions $[0, 1)$. Then, if $(t, u) \in U_v$ and $u \in I_t^-(v, i, j)$, word v_i is spoken by v_j to v , which defines the process. Using this construction and given $C, R > 0$ we obtain upper bounds $\mathbf{C}_t(w), \mathbf{R}_t(w)$ on $\mathcal{C}_t(w), R_t(w)$ for all $w \in V$, valid up to the time

$$\begin{aligned} T_{C,R} &= \min_{w \in V} T_w(C, R) \text{ where} \\ T_w(C, R) &= \inf\{t : \sum_{v \in \mathcal{C}_t(w) \cap H_t} N_t(v)^{-1} \geq C \text{ or } \mathbf{R}_t \geq R(w)\}. \end{aligned}$$

That is, we obtain for each $w \in V$ a pair of processes $\mathbf{C}_t(w), \mathbf{R}_t(w)$ with nice properties, such that $\mathcal{C}_t(w) \subseteq \mathbf{C}_t(w)$ and $R_t(w) \leq \mathbf{R}_t(w)$ for $t \leq T_{C,R}$ pointwise on realizations of the process. Given $w \in V$, $\mathbf{C}_t(w), \mathbf{R}_t(w)$ are non-decreasing and defined as follows. For $i \in \{1, \dots, n\}$ let

$$\begin{aligned} b_t(v, i) &= \sum_{k=1}^{i-1} R_t(v_k, v), \text{ and for } x \in [0, 1) \text{ let} \\ I_t(v, i, x) &= [b_t(v, i), b_t(v, i) + x) \pmod{1}. \end{aligned}$$

Define

$$N_t^\ell(v, i, R) = |\{(s, u) \in U_v : s \leq t, u \notin I_t(v, i, R)\}| \leq N_t^\ell(v).$$

Let i be such that $w = v_i$. Initially, $\mathbf{C}_0(w) = \{w\}$ and $\mathbf{R}_0(w) = 1 + C$. $\mathbf{R}_t(w)$ is defined as follows.

$$\mathbf{R}_t(w) = 1 + C + \sum_{v \in \mathbf{C}_t(w) \setminus \{w\}} 1/(1 + N_t^\ell(v, i, R)).$$

Then, $\mathbf{C}_t(w)$ is defined as follows.

$$\begin{aligned} \text{if } (t, u) \in U_v \text{ and } u \in I_{t-}(v, i, \mathbf{R}_{t-}/(n-1)), \\ \text{then } \mathbf{C}_t = \mathbf{C}_{t-} \cup \{v\}. \end{aligned}$$

We demonstrate the claimed comparison.

Lemma 7. *For each $w \in V$, $\mathcal{C}_t(w) \subseteq \mathbf{C}_t(w)$ and $R_t(w) \leq \mathbf{R}_t(w)$ for $t < T_{C,R}$.*

Proof. Let

$$\tau_c(w) = \inf\{t : \mathcal{C}_t(w) \neq \emptyset\},$$

then $\mathcal{C}_t(w) \subset \{w\} \subseteq \mathbf{C}_t(w)$ and $R_t(w) \leq 1 \leq \mathbf{R}_t$ for $t < \tau_c(w)$. For the remainder, assume $t \geq \tau_c(w)$ and let i be such that $w = v_i$. By construction, $v \in V$ is added to $\mathcal{C}_t(w)$ if

$$v \notin \mathcal{C}_{t-}(w), (t, u) \in U_v \text{ and } u \in I_{t-}(v, i, R_t(w, v)) \quad (13)$$

and otherwise, $\mathcal{C}_t(w)$ does not increase. If $t \geq \tau_c(w)$ then $N_t(w) \geq 1$, and if $z \notin H_t$ then $N_t^\ell(z) = N_t(z)$. So, from the second line of (12),

$$\begin{aligned} (n-1)R_t(w, v) &= \sum_{z \in \mathcal{C}_t(w) \setminus \{v\}} 1/N_t(z) \\ &\leq \sum_{z \in \mathcal{C}_t(w)} 1/N_t^\ell(z) + \sum_{z \in \mathcal{C}_t(w) \cap H_t} 1/N_t(z). \end{aligned}$$

If $w \in \mathcal{C}_t(w)$ then $N_t(w)^{-1} \leq 1$. By definition of $T_{C,R}$, if $\mathcal{C}_t(w) \subseteq \mathbf{C}_t(w)$ and $t < T_{C,R}$ then

$$(n-1)R_t(w, v) \leq (1 + C + \sum_{z \in \mathcal{C}_t(w) \setminus \{w\}} 1/N_t^\ell(z)).$$

If $v \in \mathbf{C}_t(w)$ and $t < T_{C,R}$ then $N_t^\ell(v) \geq N_t^\ell(v, R) + 1$, since this implies existence of a point in

$$U_v \cap \{(s, u) : s \leq t \text{ and } u \in I_s(v, i, \mathbf{R}_t(w)/(n-1))\},$$

which is counted in $N_t^\ell(v)$ but not in $N_t^\ell(v, R)$. If $\mathcal{C}_t(w) \subseteq \mathbf{C}_t(w)$ it follows that $R_t(w, v) \leq \mathbf{R}_t(w)/(n-1)$ for each v which implies the containment $\mathcal{C}_t(w) \subseteq \mathbf{C}_t(w)$ is preserved across transitions (13) that cause $\mathcal{C}_t(w)$ to increase. Since $\mathbf{C}_t(w)$ is non-decreasing and transitions are well-ordered this implies $\mathcal{C}_t(w) \subseteq \mathbf{C}_t(w)$ for $t < T_{C,R}$. It remains to check $R_t(w) \leq \mathbf{R}_t(w)$ for $\tau_c(w) \leq t < T_{C,R}$. But in this case, (10) and the previous argument give

$$R_t(w) = \sum_{v \in \mathcal{C}_t(w)} 1/N_t(v) \leq 1 + C + \sum_{v \in \mathcal{C}_t(w) \setminus \{w\}} 1/N_t^\ell(v) \leq \mathbf{R}_t(w).$$

□

Next we fix w and examine $\mathbf{C}_t(w), \mathbf{R}_t(w)$ assuming $t < T_{C,R}$, and dropping the (w) for neatness. Notice that $|\mathbf{C}_t|$ is non-decreasing and increases by 1 at rate at least $(1+C)(n-|\mathbf{C}_t|)/(n-1)$, which implies $\lim_{t \rightarrow \infty} |\mathbf{C}_t| = n$. Since $|\mathbf{C}_t|$ increases by one at a time, let y_1, \dots, y_n be the order in which vertices are added to \mathbf{C}_t , with $w = y_1$, and condition on (y_1, \dots, y_n) . We track $Z_t = |\mathbf{C}_t|$ and $N_t^i = N_t^\ell(y_i, R)$, $i = 1, \dots, n$ which suffices to determine $\mathbf{C}_t, \mathbf{R}_t$. Let $t_i = \inf\{Z_t = i\}$ denote the time at which y_i is added to \mathbf{C}_t , and let k be such that $w = v_k$. For $i \in \{2, \dots, n\}$, t_i is the least value of t such that there is a point

$$(t, u) \in \bigcup_{j \geq i} U_{y_j} \cap \{(s, v) : s \in [t_{i-1}, \infty), v \in I_s(y_j, k, \mathbf{R}_{t_{i-1}}/(n-1))\},$$

and in addition, this point belongs to U_{y_i} . Using this and basic properties of exponential random variables, together with the thinning property of the Poisson process, we find that conditioned on (y_1, \dots, y_n) ,

$$(Z_t, N_t^1, \dots, N_t^n)_{t < T_{C,R}}$$

is a Markov chain with the following transitions:

$$\begin{aligned} Z_t &\rightarrow Z_t + 1 && \text{at rate } \mathbf{R}_t(n - Z_t)/(n - 1), \text{ and} \\ \text{for } i = 1, \dots, n, N_t^i &\rightarrow N_t^i + 1 && \text{at rate } 1 - R/(n - 1). \end{aligned}$$

In particular, $\{(N_t^i)_{t < T_{C,R}} : i = 1, \dots, n\}$ is an i.i.d. collection of Poisson processes with intensity $1 - R/(n - 1)$. Since the above does not depend on the choice of values for (y_1, \dots, y_n) the same holds unconditionally. Thus Z_t can be viewed as follows: initially $Z_0 = 1$, then subject to the random environment determined by the $\{(N_t^i)\}_{i=2}^n$, Z_t increases by 1 at rate $\mathbf{R}_t(n - Z_t)/(n - 1)$. Let

$$\Lambda_t(z) = 1 + C + \sum_{i=2}^z 1/(1 + N_t^i)$$

and let (X_t) denote the process with $X_0 = 1$ that increases by 1 at rate $\Lambda_t(X_t)$. Since $(n - Z_t)/(n - 1) \leq 1$ and Λ_t is non-decreasing in z , it follows that

$$(Z_t, \mathbf{R}_t)_{t < T_{C,R}} \text{ is dominated by } (X_t, \Lambda_t(X_t)). \quad (14)$$

We can think of (X_t) as a branching process with immigration rate $1 + C$, in which individual i produces offspring at the time-decreasing rate $1/(1 + N_t^i)$. Two tasks lie ahead. The first is to estimate (X_t) . The second is to estimate $T_{C,R}$. We then combine the results to obtain Theorem 3. This is outlined as follows.

Proposition 3. *Let $b = 1 + C$. For small $\epsilon > 0$, $b \leq (8 \log n)^4$ and $R = o(n)$,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\sup_{t \leq T_{C,R}} S_t/(1+t)^{1+\epsilon} > (\log n)^9) = 0.$$

Proposition 4. *For small $\epsilon > 0$, $b = (8 \log n)^4$ and $R = b + (\log n)^{11}$,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(T_{C,R} \leq n^{1/2-\epsilon}) = 0.$$

Proof of Theorem 3. Use Propositions 3 and 4 with $b = (8 \log n)^4$ and $R = b + (\log n)^{11}$. □

4.2.1 Estimation of (X_t)

Since n does not appear in the definition of (X_t) we may as well define it using an infinite sequence $\{(N_t^i)_{t \geq 0} : i = 1, 2, \dots\}$ of Poisson processes with intensity $r = 1 - R/(n - 1)$. Clearly $r \leq 1$. Since R will be chosen $o(n)$, we will have $r \rightarrow 1$ as $n \rightarrow \infty$, so throughout we assume $r \geq 1/2$.

We begin with a useful heuristic. Let $b = 1 + C$. Replacing N_t^i with its expectation rt , X_t increases by 1 at rate $b + X_t(1 + rt)^{-1}$, which we approximate with the differential equation

$$x' = b + x/(1 + rt).$$

Let $m(t) = \exp(\int_0^t (1+rs)^{-1} ds) = (1+rt)^{1/r}$. The above equation is linear and has solution

$$x(t) = m(t)x(0) + bm(t) \int_0^t ds/m(s).$$

If r is close to 1 then $x(t)$ grows just a bit faster than linearly in time. In order to analyze (X_t) we break it up into two steps:

1. Up to a fixed time T , when the N_t^i are fairly small.
2. From time T to ∞ , when the N_t^i are fairly large.

The reason to do this is because the estimates that say $|N_t^i - rt| = o(rt)$ are only effective once rt has had time to increase. The following is the main result of this subsection.

Proposition 5. *Let $b = 1 + C$. There exist $M, x_0 \in [1, \infty)$ so that for $r \geq 1/2$ and $x \geq b \vee x_0$,*

$$\mathbb{P}(\sup_{t \geq 0} X_t - Mx(x + \log(1+t))(1+t)^{1/r} > 0) \leq 19x^{3/4}e^{-x^{1/4}/4}.$$

Recall $S_t = \max_w |\mathcal{C}_t(w)|$. Using this result we can prove Proposition 3.

Proof of Proposition 3. For each $w \in V$, using Lemma 7 and (14),

$$(|\mathcal{C}_t(w)|)_{t \leq T_{C,R}} \text{ is dominated by } (X_t).$$

Applying the result of Proposition 5 and taking a union bound over w , if $r \geq 1/2$ and $x \geq x_0$ then

$$\mathbb{P}(\sup_{t \leq T_{C,R}} S_t - \Phi(t, x) > 0) \leq 19nx^{3/4}e^{-x^{1/4}/4}, \text{ where } \Phi(t, x) = Mx(x + \log(1+t))(1+t)^{1/r}$$

If $R = o(n)$ then recalling that $r = 1 - R/(n-1)$, $1/r \leq 1 + \epsilon/3$ for large n . Since $r \leq 1$, $1+rt \leq 1+t$, and if $\epsilon > 0$ is small then $r \geq 1/2$. Letting $x = (8 \log n)^4$, the probability is $o(1)$ and since $\log(1+t) = O((1+t)^{\epsilon/2})$, it follows that

$$\Phi(t, x) = o((\log n)^9(1+t)^{1+\epsilon})$$

uniformly in t , as $n \rightarrow \infty$. □

We tackle the proof of Proposition 5 in a couple of steps.

Step 1. We obtain a somewhat crude upper bound on (X_t) that has the virtue of being effective starting at time 0. For $i \geq 1$ let $t_i = \inf\{t : X_t = i\}$, define $N_i = N_{t_i}^i$ then define Y_t, Q_t by

$$Y_0 = 1 \quad \text{and} \quad Y_t \rightarrow Y_t + 1 \quad \text{at rate} \quad Q_t = b + \sum_{i=2}^{Y_t} 1/(1 + N_{t_i}^i). \quad (15)$$

In words, at the moment t_i an individual i is added to the process, the corresponding counting process N_t^i is stopped, so that i always contributes $(1 + N_{t_i}^i)^{-1}$ to Q_t . Since $(1 + N_{t_i}^i)^{-1} \geq (1 + N_t^i)^{-1}$ for $t \geq t_i$, (X_t) is dominated by (Y_t) . The next result controls (Y_t) .

Lemma 8. *There is $M_1 \in [1, \infty)$ so that for $a \geq 2$, $b \geq 1$ and $r \geq 1/2$,*

$$\mathbb{P}(\sup_{t \geq 0} Y_t / (1 + rt)^{1+1/r} \geq abM_1) \leq 2e^{-(a-2)/2}$$

Proof. Begin by observing that (Q_t) has the concise description

$$Q_0 = b \quad \text{and} \quad Q_t \rightarrow Q_t + \Delta_t \quad \text{at rate} \quad Q_t$$

where the increment $\Delta_t \stackrel{d}{=} (1 + \text{Poisson}(rt))^{-1}$ is independently sampled every time there is a jump. Our first task is to control the size of Q_t . We compute the drift:

$$\mu_t(Q) = \ell(t)Q_t \quad \text{with} \quad \ell(t) := \mathbb{E}[\Delta_t]$$

Let $g(t) = \exp(\int_0^t \mathbb{E}[(1 + \text{Poisson}(rs))^{-1}]ds)$. Using Lemma 24 with $b(t) = 0$ and $c = 1$, for $a \geq 2$ we find

$$\mathbb{P}(\sup_t Q_t/g(t)) \geq ab) \leq e^{-(a-2)b/4} \quad (16)$$

This translates to a bound on $(Y_t)_{t \geq 0}$ as follows. Since $\mu_t(Y) = Q_t$,

$$Y_t^m = Y_t - Y_0 - \int_0^t Q_s ds$$

Since $(Y_t)_{t \geq 0}$ has transition rate Q_t and jump size exactly 1, $\sigma^2(Y_t) = Q_t$. Taking $\lambda = 1/2$ in (4) (which satisfies $c\lambda \leq 1/2$) while noting $Y_0 = 1$,

$$\mathbb{P}(Y_t \geq 1 + a + \frac{3}{2} \int_0^t Q_s ds \quad \text{for some } t \geq 0) \leq e^{-a/2}$$

Combining with (16) and taking a union bound,

$$\mathbb{P}(Y_t \geq 1 + a \left(1 + \frac{3b}{2} \int_0^t g(s)ds\right) \quad \text{for some } t \geq 0) \leq e^{-a/2} + e^{-(a-2)b/4}. \quad (17)$$

Intuitively, $g(t)$ grows roughly like $m(t)$. Let $\xi = \text{Poisson}(\lambda)$. Since $x \mapsto (1+x)^{-1}$ is convex, the inequality $\mathbb{E}[(1+\xi)^{-1}] \geq (1+\mathbb{E}[\xi])^{-1}$ goes in the wrong direction for an upper bound on $g(t)$. Anticipating our needs, we let $x = \lambda^\alpha/2$ in (35) to find

$$\mathbb{P}(\xi < \lambda - \lambda^{1/2+\alpha}/2) \leq e^{-\lambda^{2\alpha}/8} \quad \text{if } 0 < \alpha \leq 1/2. \quad (18)$$

Using the fact that $(1+\xi)^{-1} \leq 1$ and that probabilities are at most 1, then using (34),

$$\begin{aligned} \mathbb{E}[(1+\xi)^{-1}] &= \mathbb{E}[(1+\xi)^{-1} ; \xi \geq \lambda - \lambda^{1/2+\alpha}/2] + \mathbb{E}[(1+\xi)^{-1} ; \xi < \lambda - \lambda^{1/2+\alpha}/2] \\ &\leq (1 + \lambda - \lambda^{1/2+\alpha}/2)^{-1} \mathbb{P}(\xi \geq \lambda - \lambda^{1/2+\alpha}/2) + \mathbb{P}(\xi < \lambda - \lambda^{1/2+\alpha}/2) \\ &\leq (1 + \lambda - \lambda^{1/2+\alpha}/2)^{-1} + e^{-\lambda^{2\alpha}/8} \\ &\leq (1 + \lambda)^{-1} + (1 + \lambda)^{-3/2+\alpha} + e^{-\lambda^{2\alpha}/8} \end{aligned}$$

Also, if $0 < \alpha < 1/2$ and $0 < r \leq 1$ then

$$c(r, \alpha) := \int_0^\infty ((1+rs)^{-3/2+\alpha} + e^{-(rs)^{2\alpha}/8})ds < \infty.$$

Let $c(r) = \inf\{c(r, \alpha) : \alpha \in (0, 1/2)\}$ and let $c = c(1/2)$. Since $c(r, \alpha)$ decreases with r , it follows that $c(r) \leq c$ for $r \geq 1/2$. Recalling $m(t) = \exp(\int_0^t ds/(1+rs))$ defined earlier, it follows that

$$\begin{aligned} g(t) &\leq \inf_{\alpha \in (0, 1/2)} \exp\left(\int_0^t ((1+rs)^{-1} + (1+rs)^{-3/2+\alpha} + e^{-(rs)^{2\alpha}/8})ds\right) \\ &\leq e^c m(t) \end{aligned}$$

Since $m(t) = (1+rt)^{1/r}$ and $1/(r(1+1/r)) = 1/(r+1)$, and since $r > 0$, it follows that

$$\begin{aligned} \int_0^t g(s)ds &\leq e^c \int_0^t (1+rt)^{1/r} = e^c ((1+rt)^{1/r+1} - 1)/(r+1) \\ &\leq e^c ((1+rt)^{1/r+1} - 1). \end{aligned}$$

If $a \geq 2$ and $b \geq 1$ then since $c > 0$, $1 + a(1 - 3be^c/2) \leq 0$ and

$$1 + a \left(1 + \frac{3b}{2} \int_0^t g(s)ds\right) \leq \frac{3abe^c}{2} (1+rt)^{1/r+1}.$$

To conclude, take $M_1 = \frac{3}{2}e^c$, use (17) and note $e^{-a/2} + e^{-(a-2)b/4} \leq 2e^{-(a-2)/2}$ for $b \geq 2$. \square

Step 2. Next, we do two things.

1. *Lemma 9.* We control the environment $\{(N_t^i)\}_{i \geq 1}$ for $t \in [T, \infty)$.
2. *Lemma 10.* We use this to get an upper bound on (X_t) for $t \in [T, \infty)$.

Fix $\alpha \in (0, 1/2)$, then let

$$\tau_{lp}(i) = \sup\{t : N_t^i - rt + (rt)^{1/2+\alpha}/2 < 0\} \quad \text{for } i \geq 1$$

denote the last passage time of N_t^i below the curve $v(t) = rt - (rt)^{1/2+\alpha}/2$, and for $t \geq 0$ let

$$I_t = \max\{i : \tau_{lp}(j) \leq t \text{ for all } j \leq i\}.$$

Note that $\Lambda_t(x) \leq b + x/(1 + v(t))$ for $x \leq I_t$.

Lemma 9. *There is $T_0 > 0$ so that for $r \geq 1/2$ and $\alpha \in (0, 1/2)$,*

$$\mathbb{P}(\inf_{t > T} I_t - t^{1/2-\alpha} e^{(rt)^{2\alpha}/4} < 0) \leq 17T^{3/2-3\alpha} e^{-(rT)^{2\alpha}/4} \quad \text{for } T > T_0.$$

Proof. For each i , using Lemma 23 with $\lambda = r$ and $\tau_{lp}(i) = \tau_2$,

$$\mathbb{P}(\tau_{lp}(i) \geq t) \leq 4t^{1-2\alpha} e^{-(rt)^{2\alpha}/2} \quad \text{if } r \geq 1, t^{2\alpha} \geq 4.$$

Let $f(t) = 2t^{1/2-\alpha} e^{-(rt)^{2\alpha}/4}$, so the right-hand side above is $1/f(t)^2$. Then, a union bound and the fact that $f(t)^{-1} \leq 1$ gives

$$\begin{aligned} \mathbb{P}(I_t < f(t)) &= \mathbb{P}(\max_{j \leq \lceil f(t) \rceil} \tau_{lp}(j) > t) \leq \lceil f(t) \rceil / f(t)^2 \\ &\leq f(t)^{-1} (1 + f(t)^{-1}) \leq 2f(t)^{-1}. \end{aligned}$$

For $T > 0$ let

$$\begin{aligned} c_1 &= \sup_{t \geq T} f(t)/f(t+1) \quad \text{and} \\ c_2 &= (1 - 4r^{-2\alpha}(3/2 - 3\alpha)T^{-2\alpha})^{-1}. \end{aligned}$$

Note that $c_1, c_2 \in (0, \infty)$ and $\lim_{T \rightarrow \infty} c_1, c_2 = 1$ uniformly for $r \in [1/2, 1]$. Since I_t is non-decreasing, if $I_t \geq f(t)$ and $t > T$ then

$$I_{t+h} \geq f(t) \geq c_1 f(t+h) \quad \text{for } h \in [0, 1).$$

Taking a union bound over the estimate at times $T + k$, $k \geq 0$ gives

$$\mathbb{P}(\inf_{t > T} I_t - c_1 f(t) < 0) \leq \sum_{k \geq 0} 4(T+k)^{1/2-\alpha} e^{-(r(T+k))^{2\alpha}/4}.$$

The right-hand side is at most

$$4r^{-2\alpha} c_1 \int_T^\infty 4t^{1/2-\alpha} e^{-(rt)^{2\alpha}/4} dt,$$

and using (33), this is at most

$$16r^{-2\alpha} c_1 c_2 T^{3/2-3\alpha} e^{-(rT)^{2\alpha}/4}.$$

Then note that $c_1 \geq 1/2$ and $16r^{-2\alpha} c_1 c_2 \leq 17$ for T large enough, uniformly for $r \in [1/2, 1]$. \square

Lemma 10. *Given $\alpha \in (0, 1/2)$ let $\tau = \inf\{t > T : X_t > I_t\}$. There is $M_2 \in [1, \infty)$ so that for $r \in [1/2, 1]$ and $a \geq 2$,*

$$\begin{aligned} \mathbb{P}(\sup_{T \leq t < \tau} X_t - M_2(ax/(1+rT)^{1/r} + 2b \log(1+t))(1+rt)^{1/r} > 0 \mid X_T \leq x) \\ \leq e^{-(a-2)x/4(1+b(1+rT))} \end{aligned}$$

Proof. Since, as noted before, $\Lambda_t(y) \leq b+y/(1+v(t))$ for $y \leq I_t$, it follows that for $T \leq t < \tau$ and conditioned on $X_T \leq x$, (X_t) is dominated by the process \tilde{X}_t with $\tilde{X}_T = x$ that increases by 1 at rate $b + \tilde{X}_t/(1+v(t))$. We proceed as in the proof of Lemma 8. We have

$$\mu_t(\tilde{X}) = b + \ell(t)\tilde{X}_t \quad \text{with} \quad \ell(t) := 1/(1+v(t)).$$

For $a > 0$ let $E_a = \{\sup_{t \geq T} \tilde{X}_t/g(t) \geq ax + b \int_T^t ds/g(s)\}$, where $g(t) = \exp(\int_T^t \ell(s)ds)$, and let $\beta = b \int_T^\infty ds/g(s)^2$. Using Lemma 24 with $c = 1$, for $a \geq 2$ we find

$$\mathbb{P}(E_a) \leq e^{-(a-2)x/4(1+\beta)}.$$

Recall $v(t) = rt - (rt)^{1/2+\alpha}/2$. Using (34) with $\lambda = rt$,

$$\ell(t) = 1/(1+v(t)) \leq (1+rt)^{-1} + (1+rt)^{-3/2+\alpha}.$$

Let $c(r) = \int_0^\infty (1+rs)^{-3/2+\alpha} ds$, which is finite for $\alpha \in (0, 1/2)$ and $r \in (0, 1]$ and decreases with r . Let $c = c(1/2)$, so that $c(r) \leq c$ for $r \in [1/2, 1]$. Combining and noting $1+rt \leq 1+t$,

$$g(t) \leq e^c \exp\left(\int_T^t ds/(1+rs)\right) \leq e^c(1+t)^{1/r}/(1+rT)^{1/r}.$$

Using that $\ell(t) \geq 1/(1+rt)$, we obtain the complementary bound $g(t) \geq ((1+rt)/(1+rT))^{1/r}$. In this way

$$\begin{aligned} \beta = b \int_T^\infty ds/g(s)^2 &\leq (1+rT)^{2/r} \int_T^\infty (1+rs)^{-2/r} ds \\ &= b(1+rT)^{2/r} (2/r-1)^{-1} (1+rT)^{1-2/r} \leq b(1+rT). \end{aligned}$$

Using the more generous lower bound $g(t) \geq (1+rt)/(1+rT)^{1/r}$ and noting $\log(1+rT) \geq 0$ and $1/2 \leq r \leq 1$, we find

$$(1+rT)^{-1/r} b \int_T^t ds/g(s) \leq b \int_T^t ds/(1+rs) \leq 2b \log(1+rt) \leq 2b \log(1+t).$$

Let $M_2 = e^c$ and rearrange terms in the formula for E_a to complete the proof. \square

Proof of Proposition 5. We note the result of Lemma 8 applies to (X_t) since it is dominated by (Y_t) . Fix $\alpha = 1/4$ in Lemma 9 and 10. Fix $T > 0$ and let $L(t) = \frac{1}{2}e^{(rt)^{1/2}/8}$ and $x_1 = abM_1(1+rT)^{1+1/r}$. Let

$$\begin{aligned} E &= \{\sup_{t \leq T} X_t/(1+rt)^{1+1/r} \leq abM_1\}, \\ F &= \{\inf_{t \geq T} I_t - L(t) \geq 0\} \quad \text{and} \\ G &= \{\sup_{T \leq t < \tau} X_t - (aM_2x_1/(1+rT)^{1/r} + 2bM_2 \log(1+t))(1+rt)^{1/r} \leq 0\} \end{aligned}$$

be the complement of the event from, respectively, Lemma 8, 9 and 10. On E ,

$$\sup_{t \leq T} X_t/(1+rt)^{1/r} \leq abM_1(1+rT).$$

In particular, $X_T \leq x_1$, so using Lemma 10, for $b \geq 1$ and T large enough,

$$\begin{aligned} \mathbb{P}(G^c \cap E) &\leq \mathbb{P}(G^c \cap \{X_T \leq x_1\}) \leq \mathbb{P}(G^c \mid X_T \leq x_1) \\ &\leq e^{-(a-2)x_1/4(1+b(1+rT))} \leq e^{-(a-2)aM_1(1+rT)^{1/r}/5}. \end{aligned} \tag{19}$$

Using our choice of x_1 , on G we find

$$\begin{aligned} \sup_{T \leq t < \tau} X_t - (M(a, T) + 2bM_2 \log(1+t))(1+rt)^{1/r} &\leq 0 \quad \text{with} \\ M(a, T) &= a^2bM_1M_2(1+rT). \end{aligned}$$

Since $a, M_2 \geq 1$, on $E \cap G$ the inequality holds for all $t < \tau$. Taking $a = T^{1/2}$, since $r \geq 1/2$,

$$M(a, T) \leq TbM_1M_2(1+rT) \leq bM \cdot (1+rT)^2 \quad \text{with} \quad M = 2M_1M_2.$$

Using Lemmas 8 and 9, for $T > T_0$ and some $T_0 > 0$, we find $x_1 \geq 1$ and

$$\begin{aligned}\mathbb{P}(E^c) &\leq 2ee^{-T^{1/2}/2} \quad \text{and} \\ \mathbb{P}(F^c) &\leq 17T^{3/4}e^{-(rT)^{1/2}/4}.\end{aligned}$$

If T is large enough uniformly for $r \in [1/2, 1]$ the above and (19) show $\mathbb{P}(E^c), \mathbb{P}(G^c \cap E) \leq \frac{1}{2}T^{3/4}e^{-(rT)^{1/2}/4}$, so a union bound gives

$$\begin{aligned}\mathbb{P}((E \cap F \cap G)^c) &\leq \mathbb{P}(F^c) + \mathbb{P}(G^c) + \mathbb{P}(E^c) + \mathbb{P}(G^c \cap E) \leq 18T^{3/4}e^{-(rT)^{1/2}/4} \\ &\leq 19(1+rT)^{3/4}e^{-(1+rT)^{1/2}/4}.\end{aligned}$$

Choose T so that for x from the statement of the Proposition, $x = (1+rT)^2$. It suffices to check that $\tau = \infty$ on $E \cap F \cap G$. But on $E \cap F \cap G$, noting that $2M_2 \leq M$ and $r \leq 1$ on the second line,

$$\begin{aligned}I_t &\geq L(t) = t^{1/4}e^{(rt)^{1/2}/4} \quad \text{and} \\ X_t &\leq bM(1+rt)^{1/r}((1+rT)^2 + \log(1+t)) \quad \text{for } t \geq T.\end{aligned}$$

Since $\tau = \inf\{t > T : X_t > I_t\}$ and by assumption, $b \leq x = (1+rT)^2$, it suffices that

$$L(t) \geq M(1+rT)^2((1+t)^{1/r}((1+rT)^2 + \log(1+t)))$$

for $t \geq T$, which is true for T large enough, uniformly for $r \in [1/2, 1]$. \square

4.2.2 Estimation of $T_{C,R}$

Write $T_{C,R} = T_C \wedge T_R$, where

$$\begin{aligned}T_C &= \inf\{t : \max_w \sum_{v \in \mathbf{C}_t(w) \cap H_t} N_t(v)^{-1} \geq C\} \quad \text{and} \\ T_R &= \inf\{t : \max_w \mathbf{R}_t(w) \geq R\}.\end{aligned}$$

Proposition 6. *Let $b = 1 + C$. If $R \geq b + (\log n)^{11}$ and $b \leq (8 \log n)^4$ then*

$$\lim_{n \rightarrow \infty} \mathbb{P}(T_R \leq n^{1/2} \wedge T_C) = 0.$$

Proposition 7. *For each $\epsilon > 0$ and $b = (8 \log n)^4$, if $R \leq (\log n)^{12}$ then*

$$\lim_{n \rightarrow \infty} \mathbb{P}(T_C \leq n^{1/2-\epsilon} \wedge T_R) = 0.$$

Proof of Proposition 4. Notice that

$$T_{C,R} \leq t \Leftrightarrow T_C \leq t \wedge T_R \text{ or } T_R \leq t \wedge T_C,$$

then use Propositions 6 and 7 and take a union bound. \square

Next we prove Proposition 6, which is the simpler of the two.

Proof of Proposition 6. Since for $t < T_{C,R}$, each $\mathbf{R}_t(w)$ is dominated by $\Lambda_t(X_t)$, taking a union bound we find

$$\mathbb{P}(T_R \leq t \wedge T_C) \leq n\mathbb{P}(\sup_{s \leq t} \Lambda_s(X_s) \geq R).$$

For a given function $\Phi(t)$,

$$\{\sup_{s \leq t} \Lambda_s(X_s) \geq R\} \subset \{\sup_{s \geq 0} X_s - \Phi(s) > 0\} \cup \{\sup_{s \leq t} \Lambda_s(\Phi_s) \geq R\}.$$

Taking $x = (8 \log n)^4$ in Proposition 5,

$$\begin{aligned}\mathbb{P}(\sup_{s \geq 0} X_s - \Phi(s) > 0) &= o(1/n), \text{ where} \\ \Phi(s) &= M(8 \log n)^4((8 \log n)^4 + \log(1+s))(1+s)^{1/r}.\end{aligned}$$

We have the trivial bound $\Lambda_s(x) \leq b + x$, and it follows from the bound on $\ell(t)$ given in the proof of Lemma 10 that

$$\Lambda_s(x) \leq b + 2x(1 + rs)^{-1} \text{ for } x \leq I_t.$$

Let $L(s) = s^{1/4}e^{(rs)^{1/2}/4}$. Taking $T = 2(8 \log n)^2$, if n is large enough uniformly for $r \in [1/2, 1]$ then $\Phi(s) \leq L(s)$ for $s \geq T$ so using Lemma 9 and the above bounds, for $1/2 \leq r \leq 1$ and large n we find

$$\mathbb{P}(\sup_{s \geq 0} \Lambda_s(\Phi(s)) - (b + \Phi((2(8 \log n)^2) \vee (2\Phi(s)(1 + rs)^{-1})) > 0) = o(1/n)$$

if T is large enough. If $R = o(n/\log n)$, then noting $r = 1 - R/(n - 1)$ and $r \geq 1/2$ for large n , if $s \leq n^{1/2}$ then for large n ,

$$(1 + s)^{1/r}/(1 + rs) \leq 2(1 + s)^{1/r-1} \leq 2n^{2R/(n-1)} = 2e^{o(1)}$$

which approaches 2 as $n \rightarrow \infty$. It follows that

$$\sup_{s \leq n^{1/2}} 2\Phi(s)(1 + rs)^{-1} = O((\log n)^8),$$

and a similar estimate shows that $\Phi(2(8 \log n)^2) = O((\log n)^{10})$. The result follows. \square

It remains to prove Proposition 7. Define the non-decreasing spacetime set of points

$$\mathcal{A}_t(w) = \left\{ (v, s) : s \leq t \text{ and either } \begin{array}{l} v \in \mathbf{C}_{s-}(w) \text{ and agreement occurs at } (v, \cdot, s) \text{ or } (\cdot, v, s), \text{ or} \\ v \in H_s(w) \cap \mathbf{C}_s(w) \setminus \mathbf{C}_{s-}(w). \end{array} \right\}$$

To get a more workable quantity we will use the fact that

$$\sum_{v \in \mathbf{C}_t(w) \cap H_t} N_t(v)^{-1} \leq C_t(w) = \sum_{(v,s) \in \mathcal{A}_t(w)} 1/(1 + N_{t-s}^\ell(v)).$$

This way,

$$\text{if } \sup_{s \leq t} \max_w C_s(w) < C \text{ then } T_C > t. \quad (20)$$

So, to estimate T_C we control contributions to $C_t(w)$. Let $Q_t(w)$ denote the rate at which $\mathcal{A}_t(w)$ increases. Let $\{N_t^i : t \geq 0, i \geq 1\}$ be an independent collection of Poisson processes with intensity 1, let $Q, T > 0$ and let $N(t)$ be an independent Poisson process with intensity Q . Let $t_i = \inf\{t : N(t) = i\}$ and let

$$B_t = \sum_{i \leq N(t)} 1/(1 + N_{t-t_i}^i).$$

Let $T_Q(w) = \inf\{t : Q_t(w) > Q\}$ and $T_Q = \min_w T_Q(w)$. Then for any w ,

$$(C_t(w))_{t \leq T_Q} \text{ is stochastically dominated by } (B_t). \quad (21)$$

In the next lemma we control B_t .

Lemma 11. *For $T > 0$, $T_0 \geq 1$ and $Q \geq 1$,*

$$\mathbb{P}(\sup_{t \leq T} B_t > 2QT_0 + 4Q \log(2 \vee T)) \leq Q(2 + T)^2 e^{-T_0/16}.$$

Proof. We first control the value of B_T , then of B_t for $t \in [T - 1, T]$, then take a union bound to control the value over the interval $[0, T]$. Let $\tilde{N}(t) = N(T) - N(T - t)$ and fix $T_0 < T$. Using (4) with $X_t = \tilde{N}(t)$, $X_t^p = \langle X^m \rangle_t = Qt$, $c = 1$, $a = QT_0/2$, $\lambda = 1/2$ and $\bullet = +$ and noting that $c\lambda \leq 1/2$,

$$\mathbb{P}(\tilde{N}(t) \geq 2Q(t \vee T_0) \text{ for some } t \leq T) \leq e^{-QT_0/4} \leq e^{-T_0/4}$$

since $Q \geq 1$. Using the same result except with $X_t = N_t^i$, $X_t^p = \langle X^m \rangle_t = t$, $\bullet = -$, $a = T_0/4$ and $\lambda = 1/4$, and taking a union bound,

$$\mathbb{P}(N_t^i < t/2 \text{ for some } t \geq T_0 \text{ and } 2QT_0 < i \leq 2QT) \leq 2Q(T - T_0)e^{-T_0/16}.$$

Let $S = \{t \leq T : \tilde{N}(t) > \tilde{N}(t^-)\}$ be the jump times of $\tilde{N}(t)$, and label them in increasing order as $\tilde{t}_1, \tilde{t}_2, \dots, \tilde{t}_{\tilde{N}_T}$. On the complement of both events above, $\tilde{N}(T) < 2QT$ and $\tilde{t}_i > i/2Q$ for $i \geq 2QT_0$, and so $N_{\tilde{t}_i}^i \geq i/4Q$, and this gives

$$\begin{aligned} B_T &\leq 2QT_0 + \sum_{2QT_0 \leq i < 2QT} 1/(1 + i/4Q) \\ &\leq 2QT_0 + \int_{2QT_0}^{2QT} (1 + t/4Q)^{-1} dt \\ &= 2QT_0 + 4Q(\log(1 + T/2) - \log(1 + T_0/2)) \leq 2QT_0 + 4Q \log(2 \vee T). \end{aligned}$$

To see that this also bounds B_t for $t \in [T - 1, T]$, replace $N_{\tilde{t}_i}$ with $N_{\tilde{t}_{i-1}} \geq (i/2Q - 1)/2$ and use in the above to obtain the bound

$$2QT_0 + \int_{2QT_0}^{2QT} (1 + (t/2Q - 1)/2)^{-1} dt \leq 2Q(T_0 + 2(\log(1/2 + T/2) - \log(1/2 + T_0/2)))$$

which has the same upper bound, assuming $T_0 \geq 1$ so that $\log(1/2 + T_0/2) \geq 0$. The same works for $T_1 < T$ and $T_0 \geq 1$ to give

$$\mathbb{P}(B_t > 2QT_0 + 4Q \log T_1 \text{ for some } t \in [T_1 - 1, T_1]) \leq 2Q(T_1 - T_0)e^{-T_0/16} + e^{-T_0/4} \leq 2QT_1e^{-T_0/16}$$

since $2QT_0e^{-T_0/16} \geq e^{-T_0/4}$. Taking a union bound over $T_1 = 1, 2, \dots, [T], [T] + 1$ and noting $[T] \leq T$ gives the result. \square

It remains to prove the following result.

Proposition 8. *For small $\epsilon > 0$ and any $k > 0$, $Q = \log n$ and $R \leq (\log n)^k$,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(T_Q \leq n^{1/2-\epsilon} \wedge T_{C,R}) = 0.$$

Before proving it, we show how it implies Proposition 7. Use whp (with high probability) to denote an event whose probability tends to 1 as $n \rightarrow \infty$. Note that if E_1, E_2 whp then $E_1 \cap E_2$ whp.

Proof of Proposition 7. We want to show that $T_C > n^{1/2-\epsilon} \wedge T_R$ whp. Since $T_C \geq T_C \wedge T_Q$, if $T_C \wedge T_Q > n^{1/2-\epsilon} \wedge T_R$ then $T_C > n^{1/2-\epsilon} \wedge T_R$. Moreover

$$T_C \wedge T_Q > n^{1/2-\epsilon} \wedge T_R \Leftrightarrow T_C > n^{1/2-\epsilon} \wedge T_R \wedge T_Q \text{ and } T_Q > n^{1/2-\epsilon} \wedge T_R \wedge T_C.$$

Proposition 8 says that $T_Q > n^{1/2-\epsilon} \wedge T_R \wedge T_C$ whp, so it is enough to show that if $b = (8 \log n)^4$ and $Q = \log n$ then $T_C > n^{1/2-\epsilon} \wedge T_R \wedge T_Q$ whp, or equivalently that

$$\mathbb{P}(T_C \leq n^{1/2-\epsilon} \wedge T_R \wedge T_Q) = o(1).$$

In Lemma 11 take $T = n^{1/2}$, $T_0 = 48 \log n$ and $Q = \log n$ to find that

$$\mathbb{P}(\sup_{t \leq n^{1/2}} B_t > 98(\log n)^2) = O(n^{-2} \log n) = o(1/n).$$

Then, using (21) and Proposition 8 and taking a union bound over the n possible values of w ,

$$\mathbb{P}(\sup_{t \leq n^{1/2-\epsilon} \wedge T_R \wedge T_Q} \max_w C_t(w) > 98(\log n)^2) = o(1).$$

The result then follows from (20) and the fact that $98(\log n)^2 < (8 \log n)^4$ for large n . \square

By taking a union bound over w and noting the probability does not depend on w , to obtain Proposition 8 it is sufficient to show that for any w and small $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\sup_{t \leq n^{1/2-\epsilon}} Q_t(w) > \log n) = o(1/n), \quad (22)$$

noting that the probability is the same for any w . There are three ways that $\mathcal{A}_t(w)$ increases:

1. a site already in H_t is added to \mathbf{C}_t ,
2. agreement occurs at a site already in \mathbf{C}_t , or
3. a site is added simultaneously to \mathbf{C}_t and H_t .

Let $Q_t^i(w)$, $i = 1, 2, 3$ denote the rate of each event, so that $Q_t(w) = \sum_{i=1}^3 Q_t^i(w)$. Since each site in $V \setminus \mathbf{C}_t(w)$ is added to $\mathbf{C}_t(w)$ at rate $\mathbf{R}_t(w)/(n-1) \leq R/(n-1)$,

$$Q_t^1(w) \leq |H_t|R/(n-1). \quad (23)$$

Since there are $|\mathbf{C}_t(w) \cap \mathcal{C}_t(v)|$ sites in $\mathbf{C}_t(w)$ that can agree on word v , and each word is spoken at rate at most $R/(n-1)$ to each site,

$$\begin{aligned} Q_t^2(w) &\leq \sum_v |\mathbf{C}_t(w) \cap \mathcal{C}_t(v)| R/(n-1) \\ &\leq \sum_{v \neq w} |\mathbf{C}_t(w) \cap \mathcal{C}_t(v)| R/(n-1) + S_t R/(n-1), \end{aligned} \quad (24)$$

recalling that $S_t = \max_w |\mathcal{C}_t(w)|$ is the size of the largest cluster. Each time a person speaks, the probability that agreement occurs is at most $S_t/(n-1)$. Since $\mathbf{C}_t(w)$ increases at rate $\leq R$, it follows that

$$Q_t^3(w) \leq S_t R/(n-1). \quad (25)$$

The reader may think that $Q_t^3(w)$ should be 0, since a new addition to a cluster does not yet know the word. However, the upper bound cluster $\mathbf{C}_t(w)$ can grow when in the process itself, a word other than w is being spoken. Using Proposition 5 we control $Q_t^1(w)$ and $Q_t^3(w)$, which is two thirds of Proposition 8.

Lemma 12. *For each w , small $\epsilon > 0$, $R \leq n^\epsilon$ and $i = 1, 3$,*

$$\mathbb{P}\left(\sup_{t \leq n^{1/2-\epsilon} \wedge T_{C,R}} Q_t^i(w) > 1\right) = o(1/n).$$

Proof. From (23) and (25) and the choice of R , it suffices to show that

$$\mathbb{P}\left(\sup_{t \leq n^{1/2-\epsilon} \wedge T_{C,R}} \max S_t, |H_t| > n^{1-\epsilon}\right) = o(1/n).$$

The result of Proposition 3 holds with the probability being $o(1/n)$ – to see this, take $x = (12 \log n)^4$ in the proof. This gives

$$\mathbb{P}\left(\sup_{t \leq T_{C,R}} S_t - \Phi(t, x) > 0\right) = o(1/n),$$

while $\Phi(t, x)$ is still $o((\log n)^9(1+t)^{1+\epsilon})$, uniformly in t as $n \rightarrow \infty$. The desired result for $i = 3$ then follows from (25), since $\sup_{t \leq n^{1/2-\epsilon}} (\log n)^9(1+t)^{1+\epsilon} = (\log n)^9(1+n^{1/2-\epsilon})^{1+\epsilon} = o(n^{1-\epsilon})$. To get the result for $i = 1$ recall from the beginning of this section that $|H_t| \leq 2A_t$, the number of agreements up to time t , and from (11) that $A_t \leq \text{Poisson}(\int_0^t S_u du)$. Using the above bound on S_t , with probability $1 - o(1/n)$,

$$\begin{aligned} \int_0^{n^{1/2-\epsilon} \wedge T_{C,R}} S_u du &\leq \int_0^{n^{1/2-\epsilon}} \Phi(u, x) du \\ &= o((\log n)^9(1+n^{1/2-\epsilon})^{2+\epsilon}) = o((\log n)^9 n^{1-3\epsilon/2-\epsilon^2}) = o(n^{1-\epsilon}) \end{aligned}$$

for large n . From Lemma 22, $\mathbb{P}(\text{Poisson}(\lambda) > 2\lambda) \leq e^{-\lambda/3}$, so it follows that

$$\mathbb{P}\left(\sup_{t \leq n^{1/2-\epsilon} \wedge T_{C,R}} |H_t| > n^{1-\epsilon}\right) \leq e^{-n^{1-\epsilon}/6} + o(1/n) = o(1/n).$$

□

It remains to control $|\mathbf{C}_t(w) \cap \mathcal{C}_t(v)|$. We'll make use of the estimates from Lemma 11, namely that for a Poisson process $N(t)$ with intensity 1,

$$\begin{aligned} \mathbb{P}(N_t \geq 2(t \vee T) \text{ for some } t \geq 0) &\leq e^{-T/4} \text{ and} \\ \mathbb{P}(N_t \leq t/2 \text{ for some } t \geq T) &\leq e^{-T/16}. \end{aligned} \quad (26)$$

First we modify slightly the construction from the beginning of Section 4.2, using a randomization trick. The reason it needs modifying is to ensure the growth of $\mathbf{C}_t(w)$ and any $\mathcal{C}_t(v)$ are not strongly correlated. Since we only randomize the location of “excess” events that expand $\mathbf{C}_t(w)$, the reader may verify that up to a random permutation of certain vertices, the marginal distribution of each $\mathbf{C}_t(w)$, and its domination of $\mathcal{C}_t(w)$, are unchanged.

To carry out the modification, make the $\{U_v\}$ doubly-augmented, that is, each U_v is again a Poisson point process with intensity 1, but on $[0, \infty) \times [0, 1]^2$ instead of $[0, \infty) \times [0, 1]$. $\mathbf{R}_t(w)$ is defined in the same way as before, and $\mathbf{C}_t(w)$ is defined as follows.

$$\begin{aligned} & \text{if } (t, u_1, u_2) \in U_v \text{ and } u_1 \in I_{t-}(v, i, R_{t-}(w)/(n-1)), \\ & \text{or if } (t, u_1, u_2) \in U_v, u_1 \notin I_{t-}(v, i, R_{t-}(w)/(n-1)) \\ & \quad \text{and } u_2 \leq (n-1)^{-1}(\mathbf{R}_{t-}(w) - R_{t-}(w))/(1 - R_{t-}(w)), \\ & \text{then } \mathbf{C}_t = \mathbf{C}_{t-} \cup \{v\}. \end{aligned}$$

In other words,

- if $\mathcal{C}_t(w)$ was about to include v , then $\mathbf{C}_t(w)$ will too, and
- if $\mathbf{C}_t(w)$ increases when $\mathcal{C}_t(w)$ does not, then with respect to what other clusters are doing, it does so as randomly as possible.

We now control the size of $\mathbf{C}_t(w) \cap \mathcal{C}_t(v)$, for any $v \neq w$. “wp” is shorthand for “with probability”.

Lemma 13. *For any $\epsilon, k > 0$, if $R \leq n^{\epsilon/4}$ then*

$$\mathbb{P}\left(\sup_{t \leq n^{1/2-\epsilon} \wedge T_{C,R}} \sum_{v \neq w} |\mathbf{C}_t(w) \cap \mathcal{C}_t(v)| \geq n/(\log n)^k\right) = o(1/n).$$

Proof. Let $K_t = \sum_{v \neq w} |\mathbf{C}_t(w) \cap \mathcal{C}_t(v)|$. There are three ways K_t can increase.

1. $\mathbf{C}_t(w)$ acquires a site that belongs to some (possibly many) $\mathcal{C}_t(v)$, $v \neq w$,
2. some $\mathcal{C}_t(v)$, $v \neq w$ acquires a site that belongs to $\mathbf{C}_t(w)$, and
3. $\mathbf{C}_t(w)$ and some $\mathcal{C}_t(v)$, $v \neq w$ simultaneously acquire the same site.

It suffices to show the contribution to $\sup_{t \leq n^{1/2-\epsilon} \wedge T_{C,R}} K_t$ from each item is $o(n/(\log n)^k)$ wp $1 - o(1/n)$. For item 1, the increase is at most $\max_v N_t^\ell(v)$, while for items 2,3 the increase is at most 1, since at each transition, $\mathcal{C}_t(v)$ increases for at most one v , and by at most 1. Let $R_1(t), R_2(t), R_3(t)$ denote the rate of each event. Then,

$$\begin{aligned} R_i(t) & \leq \mathbf{R}_t(w) \text{ for } i \in \{1, 3\}, \text{ and} \\ R_2(t) & \leq |\mathbf{C}_t(w)| \sum_{v \neq w} R_t(v)/(n-1). \end{aligned}$$

Since $N_t^\ell(v) \leq N_{n^{1/2}}^\ell(v)$ for $t \leq n^{1/2}$ and each v , and since each $N_{n^{1/2}}^\ell(v) \sim \text{Poisson}(n^{1/2})$, using (26) with $t = T = n^{1/2}$ and a union bound,

$$\mathbb{P}\left(\sup_{t \leq n^{1/2}} \max_v N_t^\ell(v) > n^{1/2+\epsilon/2}\right) \leq ne^{-n^{\epsilon/2}/4} = o(1/n).$$

For $t < T_{C,R}$, $\mathbf{R}_t(w) \leq R$, so wp $1 - o(1/n)$, the contribution from item 1 is at most $\text{Poisson}(n^{1/2+\epsilon/2} R n^{1/2-\epsilon}) = \text{Poisson}(R n^{1-\epsilon/4-\epsilon^2/2})$ which if $R \leq n^{\epsilon/4}$ is at most $2n^{1-\epsilon^2/2} = o(n/(\log n)^k)$ for any fixed k wp $1 - o(1/n)$. Similarly, but more simply since the increase per transition is 1, the contribution from item 3 is at most $\text{Poisson}(R n^{1/2-\epsilon})$ which is at most $n^{1/2} = o(n/(\log n)^k)$ wp $1 - o(1/n)$.

For item 2, note that $\sum_{v \neq w} R_t(v) \leq \sum_v R_t(v) = n$ and that for $t < T_{C,R}$, $\mathbf{C}_t(w)$ is dominated by X_t . Applying Proposition 5, bounding $\log(1+t)$ by $\log(1+n^{1/2})$ for $t \leq n^{1/2}$ and using the trivial but convenient $n/(n-1) \leq 2$ for $n \geq 2$, we find that for $x \geq 1 + C$ large enough,

$$\mathbb{P}\left(\sup_{t < n^{1/2} \wedge T_{C,R}} R_2(t) \geq 2Mx(x + \log(1 + n^{1/2}))(1+t)^{1/r}\right) \leq 19x^{3/4}e^{-x^{1/4}/4}.$$

reactants		product	$n \cdot (\text{rate})$
$A + AB$	\rightarrow	$2AB$	$1/2$
$A + AB$	\rightarrow	$2A$	$3/2$
$B + AB$	\rightarrow	$2AB$	$1/2$
$B + AB$	\rightarrow	$2B$	$3/2$
$AB + AB$	\rightarrow	$2A$	1
$AB + AB$	\rightarrow	$2B$	1
$A + B$	\rightarrow	$B + AB$	1
$A + B$	\rightarrow	$A + AB$	1

Table 1: The various types of transitions occurring between two individuals

Taking $x = (12 \log n)^4$ the probability above is $o(1/n^2)$. Thus the contribution from item 2 is at most $\text{Poisson}(f(n^{1/2-\epsilon}))$, where

$$\begin{aligned} f(t) &= \int_0^t 2M(12 \log n)^4 ((12 \log n)^4 + \log(1 + n^{1/2}))(1+s)^{1/r} ds \\ &\leq 4M(12 \log n)^8 (1+t)^{1+1/r}, \end{aligned}$$

using $\log(1 + n^{1/2}) \leq (12 \log n)^4$ and $1 + 1/r \geq 1$. If $R = o(n)$ then for any $\epsilon > 0$, $1/r \leq 1 + \epsilon$ for large n . Therefore

$$f(n^{1/2-\epsilon}) = O((\log n)^8 n^{(1/2-\epsilon)(2+\epsilon)}) = O((\log n)^8 n^{1-3\epsilon/2-\epsilon^2}) = o(n/(\log n)^k).$$

It follows as before that $\text{Poisson}(f(n^{1/2-\epsilon})) = o(n/(\log n)^k)$ wp $1 - o(1/n)$, and the proof is complete. \square

Combining this with the other term in (24) we control $Q_t^2(w)$.

Lemma 14. *For any $k > 0$ and small $\epsilon > 0$, each w and $R \leq (\log n)^k$,*

$$\mathbb{P}\left(\sup_{t \leq n^{1/2-\epsilon} \wedge T_{C,R}} Q_t^2(w) > 2\right) = o(1/n).$$

Proof. From the proof of Lemma 12 we know that $\mathbb{P}(\sup_{t \leq n^{1/2-\epsilon} \wedge T_{C,R}} S_t > n^{1-\epsilon}) = o(1/n)$. Using this, (24), $R \leq (\log n)^k$ and Lemma 13,

$$\mathbb{P}\left(\sup_{t \leq n^{1/2-\epsilon} \wedge T_{C,R}} Q_t^2(w) > (n/(\log n)^k) + n^{1-\epsilon}(\log n)^k/(n-1)\right) = o(1/n).$$

If n is large then $(n/(\log n)^k) + n^{1-\epsilon}(\log n)^k/(n-1) \leq 2$ and the result follows. \square

Proof of Proposition 8. This follows from (22), and Lemmas 12 and 14. \square

5 Final phase

5.1 Markov chain and ODE heuristic

Using the notation of chemical reactions, we describe the eight types of interactions between any pair of individuals in Table 1. Using this as a reference, we write down the eight transitions for the three coordinates of our Markov chain as well as for $u = |x - y|$ which, as we will see later, is a key quantity in our analysis in Table 2. Note that we have rescaled (X_t, Y_t, Z_t) to $(x_t, y_t, z_t) = n^{-1}(X_t, Y_t, Z_t)$. Also note that $\Delta_i(\dots)$ and q_i are respectively the change in quantity \dots and the transition rate at the i^{th} transition.

Note that $x_t + y_t + z_t = 1$ for $t \geq 0$. To get an idea of what to expect, notice that as $n \rightarrow \infty$, sample paths approach solutions to the ODE system with $z = 1 - (x + y)$ and

$$\begin{aligned} x' &= xz + z^2 - xy \\ y' &= yz + z^2 - xy \end{aligned} \tag{27}$$

$n\Delta_i x$	$n\Delta_i y$	$n\Delta_i z$	$n\Delta_i u$	$n^{-1}q_i$
-1	0	1	$-\text{sgn}(x-y) + \mathbf{1}(u=0)$	$xz/2$
1	0	-1	$\text{sgn}(x-y) + \mathbf{1}(u=0)$	$3xz/2$
0	-1	1	$\text{sgn}(x-y) + \mathbf{1}(u=0)$	$yz/2$
0	1	-1	$-\text{sgn}(x-y) + \mathbf{1}(u=0)$	$3yz/2$
2	0	-2	$2(\text{sgn}(x-y) + \mathbf{1}(n(x-y) \in \{0, -1\}))$	$z(z - n^{-1})/2$
0	2	-2	$2(-\text{sgn}(x-y) + \mathbf{1}(n(x-y) \in \{0, 1\}))$	$z(z - n^{-1})/2$
-1	0	1	$-\text{sgn}(x-y) + \mathbf{1}(u=0)$	xy
0	-1	1	$\text{sgn}(x-y) + \mathbf{1}(u=0)$	xy

Table 2: List of transitions with jumps Δ_i and rates q_i

that has the invariant set

$$\Lambda := \{(x, y) \in \mathbb{R}_+^2 : x + y \leq 1\}.$$

The subset $\ell = \{(x, y) \in \Lambda : x = y\}$ is also invariant, since if $x = y$ then $(x - y)' = (x - y)z = 0$. Adding the x' and y' equations, the dynamics on ℓ is described by

$$z' = (1/2)(1 - 4z - z^2)$$

that has the stable fixed point $z^* = -2 + \sqrt{5}$. Thus, (27) has the equilibrium point

$$\left(\frac{1 - z^*}{2}, \frac{1 - z^*}{2}\right) = \left(\frac{3 - \sqrt{5}}{2}, \frac{3 - \sqrt{5}}{2}\right)$$

whose stable manifold contains ℓ . For the dynamics off ℓ , let $u = |x - y|$ as defined above, taking values in $[0, 1]$. From (27), we derive

$$\begin{aligned} u' &= uz \\ z' &= (1/2)(1 - u^2 - 4z - z^2) \end{aligned} \tag{28}$$

We see that u is non-decreasing, so $u(\infty) := \lim_{t \rightarrow \infty} u(t)$ is well-defined. If $u(0) > 0$ then $u(\infty) > 0$, and if in addition $u(\infty) < 1$ then according to (28), $z(t)$ has a positive limit, which contradicts $\lim_{t \rightarrow \infty} u'(t) = 0$, therefore we must have $u(\infty) = 1$.

To see the connection to Theorem 2, note that if $u(0) \geq n^{-1}$, then since the eigenvalues of the linearization near $u = 0$ and $u = 1$ are both non-zero, it should take about constant times $\log n$ amount of time for u to exceed $1 - n^{-1}$. Next, we delve into the land of martingales to make this intuition precise.

5.2 Controlling sample paths

Defining the process u by $u_t = |x_t - y_t|$, we are interested in the time to consensus, that we can express as

$$\inf \{t : u_t = 1\}.$$

Using the notation from just above, the drift and diffusivity take the form

$$\mu(u) = \sum_i q_i(u) \Delta_i(u) \quad \text{and} \quad \sigma^2(u) = \sum_i q_i(u) \Delta_i^2(u).$$

We'll write for now with u but the same holds for x, y, z and other functions of the state variables. For efficiency of notation, we'll allow the function to change depending on the variable, so $\mu(u)$ is different from $\mu(x)$ and $\mu(y)$. Also, instead of the compensator u^p we'll use the *predictor* $\bar{u} = u_0 + u^p$ which includes the initial value, and we'll denote u^m by $M(u)$, and $\langle u^m \rangle$ simply by $\langle u \rangle$. Then, $M(u) = u - \bar{u}$, and \bar{u} and $\langle u \rangle$ can be written

$$\bar{u}_t = u_0 + \int_0^t \mu_s(u) ds \quad \text{and} \quad \langle u \rangle_t = \int_0^t \sigma_s^2(u) ds.$$

Define the *jump size* $c_\Delta(u) = \sup_{u,i} |\Delta_i(u)|$. From (4), if $a > 0$, $0 < c_\Delta(u)\lambda \leq 1/2$ and $\bullet \in \pm$ then

$$\mathbb{P}(\bullet(u_t - \bar{u}_t) \geq a + \lambda \langle u \rangle_t \text{ for some } t \geq 0) \leq e^{-\lambda a}.$$

Defining the maximum transition rate $c_q(u) = \sup_u \sum_i q_i(u)$, we have the basic inequality $\sigma^2(u) \leq c(u) := (c_q c_\Delta^2)(u)$ and we obtain the corollary

$$\mathbb{P}(\bullet(u_t - \bar{u}_t) \geq a + \lambda c(u)t \text{ for some } t \geq 0) \leq e^{-\lambda a}.$$

For any quantity \cdot , we always have $c_q(\cdot) \leq n$, since there are $n(n-1)$ directed edges each ringing at rate $1/(n-1)$, and for most quantities of interest, $c_\Delta(\cdot) \leq jn^{-1}$ for a smallish integer j , giving $c_q c_\Delta^2 \leq j^2 n^{-1}$, allowing us to take λ equal to a small multiple of n while still keeping $\lambda c_q c_\Delta^2 t = O(1)$. When the context is clear, we omit the variable and simply write c, c_q, c_Δ .

The workflow of estimates is as follows. For any $\alpha > 0$, we find $\epsilon > 0$ so that the following holds with probability $1 - o(1)$ as $n \rightarrow \infty$. Item numbers correspond to the Lemmas where they are proved.

1. So long as $u_t \leq 2\epsilon$, get $|z_t - z^*| < 2\epsilon$ within constant time and keep $|z_t - z^*| < 3\epsilon$ for n time.
2. So long as $|z_t - z^*| \leq 3\epsilon$, get $u_t > 2\epsilon$ within $((2z^*)^{-1} + \alpha) \log n$ time, and find initial conditions so that $u_t \leq 2\epsilon$ for at least $((2z^*)^{-1} - \alpha) \log n$ time.
3. Once $u_t > 2\epsilon$, keep $u_t \geq \epsilon$ for $n^{1/2}$ amount of time.
4. So long as $u_t < 1 - \epsilon$, get $z_t > \epsilon/4$ within constant time and keep $z_t \geq \epsilon/12$ for n time.
5. So long as $u_t \geq \epsilon$ and $z_t \geq \epsilon/12$, get $u_t > 1 - \epsilon$ within constant times ϵ^{-2} time.
6. Once $u_t > 1 - \epsilon$, keep $u_t \geq 1 - 2\epsilon$ for $n^{1/2}$ time, and show that if $u_t \leq 1 - \epsilon + 2n^{-1}$ then so long as $u_t \geq 1 - 2\epsilon$, $u_t < 1$ for at least $(1 - \alpha) \log n$ time.
7. So long as $u_t \geq 1 - 2\epsilon$, get $u_t = 1$ within $(1 + \alpha) \log n$ time.

Propositions 9–11 stitch together Lemmas 15–16, Lemmas 17–19, and Lemmas 20–21, respectively. The combination of these propositions into the proof of Theorem 2 is given at the end of this section.

Lemma 15. *Let $b_t = z_t - z^*$ and let*

$$\tau = \inf\{t : u_t \geq 2\epsilon\}, \quad \tau_0 = \tau \wedge \inf\{t : |b_t| \leq \epsilon\} \quad \text{and} \quad \tau_1 = \tau \wedge \inf\{t : |b_t| \geq 3\epsilon\}.$$

If $\epsilon \leq 1/4$ and $n \geq 8/\epsilon^2$ then

$$\mathbb{P}(\tau_0 \geq 2/\epsilon^2) \leq e^{-\epsilon^2 n/64}$$

and for integer $N > 0$,

$$\mathbb{P}(\tau_1 \leq \epsilon^2 N/6 \quad \text{and} \quad |b_{\tau_1}| \geq 3\epsilon \mid |b_0| \leq \epsilon) \leq 2Ne^{-\epsilon^4 n/8}.$$

Proof. Using Table 2, we find that

$$\mu(b) = (1/2)(-b(z+2+\sqrt{5}) - u^2) + 2zn^{-1}.$$

Since $c_\Delta(b) \leq 2n^{-1}$ and $c_q(b) \leq n$, $\sigma^2(b) \leq (c_q c_\Delta^2)(b) \leq 4n^{-1}$, and using the product rule from Lemma 4 on $b \cdot b$,

$$\begin{aligned} \mu(b^2) &= 2b\mu(b) + \sigma^2(b) \leq 2b\mu(b) + 4n^{-1} \\ &\leq b(-b(z+2+\sqrt{5}) - u^2) + 4(1+zb)n^{-1} \end{aligned}$$

If $u < 2\epsilon$ and $|b| > \epsilon$ then, since $\sqrt{5} \geq 2$ and $z \geq 0$,

$$\begin{aligned} |-b(z+2+\sqrt{5}) - u^2| &\geq |b(z+2+\sqrt{5})| - u^2 \\ &\geq 4|b| - u^2 \geq 4\epsilon - 4\epsilon^2 = 4\epsilon(1 - \epsilon) \end{aligned}$$

so if in addition $\epsilon, 2n^{-1}/\epsilon^2 \leq 1/4$, then since $zb \leq 1$,

$$\mu(b^2) \leq -4\epsilon^2(1 - \epsilon) + 8n^{-1} \leq -2\epsilon^2. \quad (29)$$

and since $b_t^2 - b_0^2 \geq -1$, we find

$$b_t^2 - \bar{b}_t^2 = b_t^2 - b_0^2 - \int_0^t \mu_s(b^2) ds \geq -1 + 2\epsilon^2 t.$$

Moreover, if $b \in [0, 1]$ then for $\delta \in \mathbb{R}$, $|(b + \delta)^2 - b^2| = |2\delta b + \delta^2| \leq 2|\delta| + \delta^2$, which implies that $c_\Delta(b^2) \leq 2(2n^{-1}) + (2n^{-1})^2 \leq 8n^{-1}$. So, we can take $\lambda \in (0, n/16)$ and $c = c_q c_\Delta^2 = 64n^{-1}$. Choosing $\lambda = \epsilon^2 n/64$ gives $c\lambda t = \epsilon^2 t$. If $\tau_0 > T$ we find

$$b_T^2 - \bar{b}_T^2 - c\lambda T \geq -1 + \epsilon^2 T.$$

Taking $T = 2/\epsilon^2$ gives a lower bound of 1. So, taking $a = 1$ and $\bullet = +$ gives the first statement. Next, let $\tau_2 = \inf\{t : |b_t - 2\epsilon| \geq \epsilon \text{ or } u_t \geq 2\epsilon\}$. Using (29), $\mu(b_t^2) \leq -2\epsilon^2$ for $t < \tau_2$. Thus, if $|b_0 - 2\epsilon| \leq n^{-1}$, $b_{\tau_2} \geq 3\epsilon$ then

$$\begin{aligned} b_{\tau_2}^2 - \bar{b}_{\tau_2}^2 &\geq (3\epsilon)^2 - (2\epsilon + n^{-1})^2 + 2\epsilon^2 \tau_2 \\ &\geq 5\epsilon^2 - n^{-1}(4\epsilon + n^{-1}) + 2\epsilon^2 \tau_2 \end{aligned}$$

Taking $\lambda = \epsilon^2 n/32$, $c\lambda t = 2\epsilon^2 t$. If $n \geq 5/\epsilon$ then $b_{\tau_2}^2 - \bar{b}_{\tau_2}^2 - c\lambda \tau_2 \geq 4\epsilon^2$. Noting that $\mathbb{P}(\tau_2 < \infty) = 1$, then taking $a = 4\epsilon^2$ and $\bullet = +$,

$$\mathbb{P}(|b_{\tau_2}| \geq 3\epsilon \mid |b_0 - 2\epsilon| \leq n^{-1}) \leq e^{-\epsilon^4 n/8} \quad (30)$$

On the other hand, since $z + 2 + \sqrt{5} \leq 6$ and $|b_t| \leq 1$, $\mu(b_t) \geq -7/2 - 6n^{-1}$. If $|b_0 - 2\epsilon| \leq n^{-1}$, $\tau_2 \leq T$ and $b_{\tau_2} \leq \epsilon$ then

$$\begin{aligned} -M_{\tau_2}(b^2) &\geq (2\epsilon - n^{-1})^2 - \epsilon^2 - (7/2 + 6n^{-1})T \\ &\geq 3\epsilon^2 - 4\epsilon n^{-1} - (7/2 + 6n^{-1})T \end{aligned}$$

If $n \geq 4/\epsilon$ then $4\epsilon n^{-1} \leq \epsilon^2$. Taking $\lambda = n/64$, $c\lambda = 1$. If $n \geq 6$ then $(7/2) + 6n^{-1} \leq 5$. Letting $T = \epsilon^2/6$,

$$-M_{\tau_2}(b^2) - c\lambda \tau_2 \geq 2\epsilon^2 - 6T \geq \epsilon^2,$$

and taking $a = \epsilon^2$ and $\bullet = -$, it follows that

$$\mathbb{P}(\tau_2 \leq \epsilon^2/6 \text{ and } |b_{\tau_2}| \leq \epsilon \mid |b_0 - 2\epsilon| \leq n^{-1}) \leq e^{-\epsilon^2 n/64} \quad (31)$$

The result follows by stopping the process each time $|b - 2\epsilon| \leq n^{-1}$, using (30) and (31), then using the Markov property and taking a union bound while noting that $e^{-\epsilon^2 n/64} \leq e^{-\epsilon^4 n/8}$ for $\epsilon \leq 1/2\sqrt{2} \leq 1/4$. \square

Lemma 16. *As in Lemma 15, let $b_t = z_t - z^*$,*

$$\tau = \inf\{t : u_t \geq 2\epsilon\} \quad \text{and} \quad \tau_1 = \tau \wedge \inf\{t : |b_t| \geq 3\epsilon\}.$$

Let

$$c_1 = z^* - 3\epsilon \quad \text{and} \quad c_2 = z^* + 3\epsilon.$$

For $C > 0$, if $\epsilon \leq \min(C^{1/2}/6, 1/60)$ and $n \geq \max(20, 64(1 + C^{1/2})^2)$ then

$$\mathbb{P}(\tau_1 > 25(3 + C) + (2c_1)^{-1} \log n) \leq e^{-n^{1/2}/1600(1+C^{1/2})^2} + e^{-c_1 C/18}$$

Also, for $C_2 > 0$,

$$\begin{aligned} \mathbb{P}(\tau_1 \leq (2c_2)^{-1}(\log n - C_2) \quad \text{and} \quad u_{\tau_1} \geq 2\epsilon \mid u_0 = C_2 n^{-1/2}) \\ \leq e^{-C_2 + 2/c_1 + n^{-1/2}} + (2\epsilon)^{-1} C_2 e^{-C_2/2} \end{aligned}$$

Proof. Notice that $\mu(u) = uz + \rho_1 \mathbf{1}(u = 0) + \rho_2 \mathbf{1}(u \leq n^{-1})$ where $0 \leq \rho_1, \rho_2 \leq 1$. Using transitions 5 and 6 from Table 2,

$$\mu(u^2) = 2u\mu(u) + \sum_i (q_i \cdot \Delta_i^2)(u) \geq 2u^2z + 4n^{-1}z(z - n^{-1})$$

If $|z - z^*| < 3\epsilon$ and $3\epsilon, n^{-1} \leq 1/20$, then since $z^* > 1/5$, it follows that $z - n^{-1} > 1/10$ and $z(z - n^{-1}) > 1/100$, so that $\mu(u^2) \geq n^{-1}/25$.

For $C > 0$, let $\tau_3 = \tau_1 \wedge \inf\{t : u^2 \geq Cn^{-1}\}$. If $u < C^{1/2}n^{-1/2}$ then since $c_\Delta(u) \leq 2n^{-1}$,

$$\begin{aligned} |\Delta_i(u^2)| &\leq 2u|\Delta_i(u)| + |\Delta_i(u)|^2 \\ &\leq 4C^{1/2}n^{-3/2} + 4n^{-2} \leq C_1n^{-3/2} \quad \text{with } C_1 = 4(1 + C^{1/2}) \end{aligned}$$

so we can take $\lambda \in (0, n^{3/2}/2C_1)$ and $c = C_1^2n^{-5/2}$. From the bound on the jump size, we find that $u_{\tau_3}^2 \leq Cn^{-1} + C_1n^{-3/2}$. If $\tau_3 > T$ then noting $u_0^2 \geq 0$ it follows that

$$-M_T(u^2) \geq -Cn^{-1} - C_1n^{-3/2} + n^{-1}T/25$$

Taking $T = 25(3 + C)$ and $\lambda = C_1^{-2}n^{3/2}/50$, $c\lambda T = 2n^{-1}$ and $n^{-1}T/25 = (3 + C)n^{-1}$, so

$$-M_T(u^2) - c\lambda T \geq n^{-1} - C_1n^{-3/2}.$$

If $n \geq (2C_1)^2$ this is at least $n^{-1}/2$, so taking $a = n^{-1}/2$ and $\bullet = -$ gives the estimate

$$\mathbb{P}(\tau_3 > 25(3 + C)) \leq e^{-C_1^{-2}n^{1/2}/100}$$

Next, let $c_1 = z^* - 3\epsilon$ and define h by $h_t = e^{-c_1 t}u_t$ so that $\mu_t(h) \geq 0$ and $-M_t(h) \geq u_0 - h_t$ for $t < \tau_1$. Here we take $c = c(t)$ that depends on time, such that $\sigma_t^2(h) \leq c(t)$, and use the more general inequality $\langle h \rangle_t \leq \int_0^t c(s)ds$. Since $|\Delta_i(u)| \leq 2n^{-1}$ for any u, i , $|\Delta_i(h_t)| \leq 2e^{-c_1 t}n^{-1} \leq 2n^{-1}$, so we can take $\lambda \in (0, n/4)$ and $c(t) = 4n^{-1}e^{-2c_1 t}$. If $u_0 \geq C^{1/2}n^{-1/2}$ and $\tau_1 > T$ then

$$-M_T(h) - \lambda \int_0^T c(t)dt \geq C^{1/2}n^{-1/2} - 2e^{-c_1 T}\epsilon - 4n^{-1}\lambda \int_0^T e^{-2c_1 t}dt.$$

Letting $T = (2c_1)^{-1} \log n$ and bounding the integral by $1/2c_1$ we obtain

$$-M_T(h) - \lambda \int_0^T c(t)dt \geq C^{1/2}n^{-1/2} - 2\epsilon n^{-1/2} - 4n^{-1}\lambda/2c_1.$$

If $C \leq (3/2c_1)^2 n$, then taking $\lambda = C^{1/2}n^{1/2}c_1/6$ and $\epsilon \leq C^{1/2}/6$, and taking $a = C^{1/2}n^{-1/2}/3$ and $\bullet = -$ we find that

$$\mathbb{P}(\tau_1 > (2c_1)^{-1} \log n \mid u_0 \geq C^{1/2}n^{-1/2}) \leq e^{-c_1 C/18}$$

To get a matching lower bound on τ_1 we need an upper bound on $\mu(u)$. If $u > n^{-1}$ and $|z - z^*| \leq 3\epsilon$ then letting $c_2 = z^* + 3\epsilon$, $\mu(u) \leq c_2 u$. Let $\tau_4 = \tau_1 \wedge \inf\{t : u_t \leq n^{-1}\}$. If $u_0 = C_2 n^{-1/2}$ for $C_2 > 0$ and $u_{\tau_4} \leq n^{-1}$ then as before, for $\lambda \leq n/4$ we find

$$-M_{\tau_4}(h) - \lambda \int_0^{\tau_4} c(t)dt \geq C_2 n^{-1/2} - n^{-1} - 2\lambda n^{-1}/c_1$$

If $u_0 > n^{-1}$ then $\tau_4 \neq \tau_1$ if and only if $u_{\tau_4} \leq n^{-1}$. Taking $\lambda = n^{1/2}$, $a = C_2 n^{-1/2} - n^{-1} - 2n^{-1/2}/c_1$ and using $\bullet = -$ we find that

$$\mathbb{P}(\tau_1 \neq \tau_4 \mid u_0 = C_2 n^{-1/2}) \leq e^{-C_2 + 2/c_1 + n^{-1/2}}$$

On the other hand, $\mu(u_t) \leq c_2 u_t$ for $t < \tau_4$, so $s_t = e^{-c_2(t \wedge \tau_4)} u_{t \wedge \tau_4}$ is a supermartingale. Using non-negativity of s_t , the fact that $t \mapsto e^{-c_2 t}$ is non-increasing, and optional stopping,

$$\begin{aligned} \mathbb{E}[u_{\tau_4}; \tau_4 \leq T] &= e^{c_2 T} \mathbb{E}[e^{-c_2 T} u_{\tau_4}; \tau_4 \leq T] \\ &\leq e^{c_2 T} \mathbb{E}[e^{-c_2 \tau_4} u_{\tau_4}] \leq e^{c_2 T} \mathbb{E}[u_0] \end{aligned}$$

Using Markov's inequality,

$$\mathbb{P}(\tau_4 \leq T \text{ and } u_{\tau_4} \geq 2\epsilon \mid u_0 = C_2 n^{-1/2}) \leq (2\epsilon)^{-1} e^{c_2 T} C_2 n^{-1/2}$$

Letting $T = (2c_2)^{-1}(\log n - C_2)$, this is at most $(2\epsilon)^{-1} C_2 e^{-C_2/2}$. The second statement then follows from a union bound. \square

Proposition 9. *Let $\tau = \inf\{t : u_t \geq 2\epsilon\}$ as in Lemma 15,16. Then for any $\alpha > 0$, there is $\epsilon_0 > 0$ so that for $\epsilon \in (0, \epsilon_0]$,*

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_{(u,z)} \mathbb{P}_{(u,z)}(\tau > ((2z^*)^{-1} + \alpha) \log n) &= 0 \\ \lim_{n \rightarrow \infty} \sup_{(u,z)} \mathbb{P}_{(u,z)}(\tau > ((2z^*)^{-1} - \alpha) \log n) &= 1 \end{aligned}$$

Proof. Given $\alpha > 0$, let $0 < \epsilon_0 \leq 1/60$ be small enough that

$$(2(z^* - 3\epsilon_0))^{-1} \leq (2z^*)^{-1} + \alpha/2 \quad \text{and} \quad (2(z^* + 3\epsilon_0))^{-1} \geq (2z^*)^{-1} - \alpha/2.$$

Recall that $b_t = z_t - z^*$ and let $\tau_0 = \tau \wedge \inf\{t : |b_t| \leq \epsilon\}$ as in Lemma 15. Fix $\epsilon \in (0, \epsilon_0]$. Using the first result of Lemma 15,

$$\mathbb{P}(\tau_0 < 2/\epsilon^2) = 1 - o(1)$$

If $\tau = \tau_0$ and $\tau_0 < 2/\epsilon^2$ then in particular, $\tau \leq ((2z^*)^{-1} + \alpha) \log n$ for large enough n . If $\tau > \tau_0$ then $|b_{\tau_0}| < \epsilon$. Letting $N = n$ in the second result of Lemma 15 and using the strong Markov property,

$$\mathbb{P}(|b_{\tau_0+t}| < 3\epsilon \text{ for all } t \leq (\tau - \tau_0) \wedge (\epsilon^2 n/6) \mid |b_{\tau_0}| \leq \epsilon) = 1 - o(1)$$

Then, letting $C = (\alpha/100) \log n$ in the first result of Lemma 16 and using again the strong Markov property,

$$\mathbb{P}(|b_{\tau_0+t}| < 3\epsilon \text{ and } u_{\tau_0+t} < 2\epsilon \text{ for all } t \leq 75 + (\alpha/4) \log n + ((2z^*)^{-1} + \alpha/2) \log n) = o(1)$$

If n is large enough then $\epsilon^2 n/6 > 75 + (\alpha/4) \log n + ((2z^*)^{-1} + \alpha/2) \log n$. Combining these results, we find that from any initial distribution, if n is large enough then

$$\mathbb{P}(\tau \geq 2/\epsilon^2 + 75 + (\alpha/4) \log n + ((2z^*)^{-1} + \alpha/2) \log n) = o(1)$$

If n is large enough then $2/\epsilon^2 + 75 \leq (\alpha/4) \log n$ and the first statement follows. For the second statement, recall that $\tau_1 = \tau \wedge \inf\{t : |b_t| \geq 3\epsilon\}$, and let $C_2 = (\alpha/4c_2) \log n$ to find that

$$\mathbb{P}(\tau_1 \leq ((2c_2)^{-1} - \alpha/2) \log n \text{ and } u_{\tau_1} \geq 2\epsilon \mid u_0 = C_2 n^{-1/2}) = o(1)$$

By definition, either $u_{\tau_1} \geq 2\epsilon$ or $|b_{\tau_1}| \geq 3\epsilon$. Combining with the second result of Lemma 15, if $\epsilon^2 n/6 > ((2c_1)^{-1} - \alpha/2) \log n$ then

$$\mathbb{P}(\tau \leq ((2c_2)^{-1} - \alpha/2) \log n \mid u_0 = C_2 n^{-1/2} \text{ and } |b_0| < \epsilon) = o(1)$$

and the second statement follows. \square

Next we show that if $u_0 \geq 2\epsilon$ then there is a good chance $u_t \geq \epsilon$ for as long as we need.

Lemma 17. *If $n > 1/\epsilon$ then for $T > 0$,*

$$\mathbb{P}(\inf_{t \leq T} u_t < \epsilon \mid u_0 \geq 2\epsilon) \leq e^{-\epsilon^2 n/16T}$$

Proof. We know that if $u > n^{-1}$ then $\mu(u) = uz \geq 0$. Since $c_\Delta(u) \leq 2n^{-1}$, we can take $\lambda \in (0, n/4)$ and $c = 4n^{-1}$. Let $\tau_0 = \inf\{t : u_t < \epsilon\}$. If $u_0 \geq 2\epsilon$ and $\tau_0 < T$ then $M_{\tau_0}(u) - c\lambda\tau_0 \geq \epsilon - 4\lambda n^{-1}T$, so using $\bullet = +$,

$$\mathbb{P}(\tau_0 < T \mid u_0 \geq 2\epsilon) \leq e^{-\lambda(\epsilon - 4\lambda n^{-1}T)}$$

Optimizing in λ then gives the result. \square

Lemma 18. Let $\tau = \inf\{t : u_t < \epsilon \text{ or } u_t > 1 - \epsilon\}$ and let

$$\tau_2 = \tau \wedge \inf\{t : z_t > \epsilon/4\} \quad \text{and} \quad \tau_3 = \tau \wedge \inf\{t : z_t < \epsilon/12\}.$$

If $\epsilon \leq 1/4$ and $n \geq 1/\epsilon$ then

$$\mathbb{P}(\tau_2 > 20) \leq e^{-\epsilon^2 n/128},$$

and for integer $N > 0$,

$$\mathbb{P}(\tau_3 \leq \epsilon N/48 \quad \text{and} \quad z_{\tau_3} < \epsilon/12 \mid z_0 > \epsilon/4) \leq 2Ne^{-n\epsilon^2/192+1/4}.$$

Proof. Since $c_\Delta(z) \leq 2n^{-1}$, we can take $\lambda \in (0, n/4)$ and $c = 4n^{-1}$. Recall that

$$\mu(z) = (1/2)(1 - u^2 - 4z - z^2) + 2zn^{-1}$$

so if $u \leq 1 - \epsilon$, $z \leq \epsilon/4$ and $\epsilon \leq 1/4$ then $\mu(z) \geq \epsilon/4$. Minding the jump size, $z_{\tau_2} \leq \epsilon/4 + 2n^{-1}$. If $\tau_2 > T$ and $\epsilon/4 \geq 4n^{-1}\lambda$ then

$$\begin{aligned} -M_{\tau_2}(z) - c\lambda\tau_2 &\geq -\epsilon/4 - 2n^{-1} + (\epsilon/4)T - 4n^{-1}\lambda T \\ &= \epsilon(T(1 - 16\lambda(\epsilon n)^{-1}) - 1 - 8(\epsilon n)^{-1})/4 \end{aligned}$$

Taking $T = 20$ and $\lambda = \epsilon n/32$, if $n \geq 1/\epsilon$ we have the lower bound $\epsilon/4$. Taking $a = \epsilon/4$ and $\bullet = -$ gives the first statement. Now, let $\tau_4 = \tau \wedge \inf\{t : |z_t - \epsilon/6| \geq \epsilon/12\}$. Suppose that $|z_0 - \epsilon/6| \leq n^{-1}$, then $\mu(z_t) \geq \epsilon/4$ for $t < \tau_4$. If $z_{\tau_4} \leq \epsilon/12$ then

$$-M_{\tau_4}(z) - c\lambda\tau_4 \geq \epsilon/6 - n^{-1} - \epsilon/12 + (\epsilon/4 - 4\lambda n^{-1})\tau_4.$$

Taking $\lambda = \epsilon n/16$ and $a = \epsilon/12 - n^{-1}$, $\mathbb{P}(z_{\tau_4} \leq \epsilon/12) \leq e^{-\epsilon^2 n/192+\epsilon/16}$. On the other hand, if $z_{\tau_4} \geq \epsilon/4$ and $\tau_4 \leq T$ then since $\mu(z) \leq 1/2 + 2n^{-1}$,

$$M_{\tau_4}(z) - c\lambda\tau_4 \geq \epsilon/12 - n^{-1} - T(1/2 + 2n^{-1} + 4\lambda n^{-1})$$

Taking $\lambda = n/4$, $T = \epsilon/48$ and $a = \epsilon/24 - n^{-1}$, if $n \geq 4$ then

$$\mathbb{P}(\tau_4 \leq \epsilon/48 \quad \text{and} \quad z_{\tau_4} \geq \epsilon/4) \leq e^{-\epsilon n/96+1/4}.$$

Combining these,

$$\mathbb{P}(\tau_4 \leq \epsilon/48 \quad \text{or} \quad z_{\tau_4} \leq \epsilon/12) \leq e^{1/4}(e^{-\epsilon^2 n/192} + e^{-\epsilon n/96})$$

The result follows by stopping the process each time $|z - \epsilon/6| \leq n^{-1}$, using the strong Markov property, taking a union bound, and using $e^{-\epsilon n/92} \leq e^{-\epsilon^2 n/192}$. \square

Lemma 19. Let $\tau = \inf\{t : u_t < \epsilon \text{ or } u_t > 1 - \epsilon\}$ as in Lemma 18 and let $\tau_5 = \tau \wedge \inf\{t : z_t < \epsilon/12\}$. Then,

$$\mathbb{P}(\tau_5 > 48/\epsilon^2) \leq e^{-\epsilon^2 n/96}$$

Proof. If $u \geq \epsilon$ and $z \geq \epsilon/12$ then $\mu(u) \geq \epsilon^2/12$. Since $c_\Delta(u) \leq 2n^{-1}$, we can take $\lambda \leq n/4$ and $c = 4n^{-1}$. If $\tau_5 > T$ then since $u_T - u_0 \leq 1$, if $\epsilon^2/12 \geq 4\lambda n^{-1}$ then

$$-M_{\tau_5}(u) - c\lambda\tau_5 \geq -1 + T(\epsilon^2/12 - 4\lambda n^{-1})$$

Taking $\lambda = -\epsilon^2 n/96$, $T = 48/\epsilon^2$, $a = 1$ and $\bullet = -$, the result follows. \square

Proposition 10. Let $\tau = \inf\{t : u_t \leq \epsilon \text{ or } u_t \geq 1 - \epsilon\}$, as in Lemma 18. Then for any $\alpha > 0$ and $\epsilon \in (0, 1/4)$,

$$\lim_{n \rightarrow \infty} \sup_{(u,z): u \geq 2\epsilon} \mathbb{P}_{(u,z)}(u_\tau \leq \epsilon \text{ or } \tau > \alpha \log n) = 0$$

Proof. Taking $T = n^{1/2}$ in Lemma 17,

$$\mathbb{P}(\inf_{t \leq n^{1/2}} u_t \leq \epsilon \mid u_0 \geq 2\epsilon) = o(1)$$

Let $\tau_2 = \tau \wedge \inf\{t: z_t > \epsilon/4\}$ as in Lemma 18. Using the first result of Lemma 18,

$$\mathbb{P}(\tau_2 \leq 20) = 1 - o(1)$$

Letting $N = n$ in the second result of Lemma 18 and using the strong Markov property,

$$\mathbb{P}(z_{\tau_2+t} \geq \epsilon/12 \text{ for all } t \leq (\tau - \tau_2) \wedge (\epsilon n/48) \mid z_{\tau_2} > \epsilon/4) = 1 - o(1)$$

Then, using the strong Markov property and the result of Lemma 19,

$$\mathbb{P}(z_{\tau_2+t} \geq \epsilon/12 \text{ and } u_{\tau_2+t} \in [\epsilon, 1 - \epsilon] \text{ for all } t \leq 48/\epsilon^2) = o(1)$$

If n is large enough then $\min(n^{1/2}, \epsilon n/48, \alpha \log n) \geq 20 + 48/\epsilon^2$. Combining the estimates gives the result. \square

Lemma 20. Let $v = \max(x, y)$, $w = \min(x, y)$ and define

$$\tau = \inf\{t: w_t = z_t = 0 \text{ or } 2w_t + z_t \geq 2\epsilon\}$$

For any $T > 0$ and $\epsilon \leq 1/4$,

$$\mathbb{P}(2w_\tau + z_\tau \geq 2\epsilon \text{ and } \tau \leq T \mid 2w_0 + z_0 \leq \epsilon) \leq e^{-\epsilon^2 n/16T}$$

Also, if $n \geq 4/\epsilon$ then for $c > 0$,

$$\mathbb{P}(w_\tau = z_\tau = 0 \text{ and } \tau \leq C \log n \mid 2w_0 + z_0 \geq \epsilon - 2n^{-1}) \leq 12\epsilon^{-1}n^{-1+(1+13\epsilon/2)C}$$

Proof. We have $v = u + w$ and $2w + z = 1 - u$. Recall that if $u > n^{-1}$ then $\mu(u) = uz$, so $\mu(2w + z) = (w - v)z$. Since $w \leq v$, $\mu(2w + z) \leq 0$. Since $c_\Delta(2w + z) \leq 2n^{-1}$, we can take $\lambda \leq n/4$ and $c = 4n^{-1}$. If $2w_0 + z_0 \leq \epsilon$, $2w_\tau + z_\tau \geq 2\epsilon$ and $\tau \leq T$ then $M_\tau(2w + z) - c\lambda\tau \geq \epsilon - 4\lambda n^{-1}T$. Taking $\lambda = n/8T$ and $\bullet = +$, the first statement follows. On the other hand, we have always $v \leq 1$, and if $t < \tau$ then $w_t \leq \epsilon$, $z_t \leq 2\epsilon$ and $v_t \geq 1 - 2\epsilon$. Looking to Table 2, ignoring the 5th and 8th transitions, ignoring some increases, and bounding the rates in the right direction it is easy to check that for $t \leq \tau$, (w_t, z_t) dominates the process $(\tilde{w}_t, \tilde{z}_t)$ with initial value (w_0, z_0) and the following transitions:

$n\Delta_i(\tilde{w})$	-1	0	0	0	0	-1
$n\Delta_i(\tilde{z})$	0	-1	1	-1	-2	0
$n^{-1}q_i$	$\tilde{w}\epsilon$	$3\epsilon\tilde{z}/2$	$(1 - 2\epsilon)\tilde{z}/2$	$3\tilde{z}/2$	$2\epsilon\tilde{z}$	\tilde{w}

Note the transition rates are linear. We easily compute

$$\mu(\tilde{w}) = -(1 + \epsilon)\tilde{w} \text{ and } \mu(\tilde{w}^2) = (1 + \epsilon)(-2\tilde{w}^2 + n^{-1}\tilde{w})$$

so that if w_0 is deterministic, we solve to obtain

$$\mathbb{E}[\tilde{w}_t] = e^{-(1+\epsilon)t}w_0 \text{ and } \mathbb{E}[\tilde{w}_t^2] = e^{-2(1+\epsilon)t}w_0^2 + n^{-1}w_0e^{-(1+\epsilon)t}(1 - e^{-(1+\epsilon)t})$$

Combining, $\text{Var}(\tilde{w}_t) \leq n^{-1}w_0e^{-(1+\epsilon)t}$ and so

$$\mathbb{P}(\tilde{w}_t = 0) \leq \mathbb{P}(|\tilde{w}_t - \mathbb{E}[\tilde{w}_t]| \geq \mathbb{E}[\tilde{w}_t]) \leq \frac{\text{Var}(\tilde{w}_t)}{(\mathbb{E}[\tilde{w}_t])^2} \leq (w_0n)^{-1}e^{(1+\epsilon)t}$$

Similarly,

$$\mu(\tilde{z}) = -(1 + 13\epsilon/2)\tilde{z} \text{ and } \mu(\tilde{z}^2) = -2(1 + 13\epsilon/2)\tilde{z}^2 + n^{-1}(2 + 17\epsilon/2)\tilde{z}$$

so that if z_0 is deterministic,

$$\mathbb{E}[\tilde{z}_t] \leq e^{-(1+13\epsilon/2)t} z_0 \quad \text{and} \quad \text{Var}(z_t) \leq n^{-1} z_0 \frac{2+17\epsilon/2}{1+13\epsilon/2} e^{-(1+13\epsilon/2)t}$$

and since the above fraction is at most 2,

$$\mathbb{P}(\tilde{z}_t = 0) \leq 2(z_0 n)^{-1} e^{(1+13\epsilon/2)t}$$

If $2w_0 + z_0 \geq a$ then $\max(w_0, z_0) \geq a/3$, so for $T > 0$

$$\begin{aligned} & \mathbb{P}(\sup_{t \leq T} (2w_t + z_t) \leq 2\epsilon, \ z_T = w_T = 0 \mid 2w_0 + z_0 \geq a) \\ & \leq \max(\mathbb{P}(\tilde{w}_T = 0 \mid w_0 \geq a/3), \mathbb{P}(\tilde{z}_T = 0 \mid z_0 \geq a/3)) \end{aligned}$$

Letting $a = \epsilon - 2n^{-1}$ and $T = C \log n$, if $n \geq 4/\epsilon$ then $a \geq \epsilon/2$ and the second statement follows. \square

Lemma 21. *Let v, w and τ be as in Lemma 20. If $\epsilon < 1/6$, $n \geq 1/\epsilon$ and $C > 0$ then*

$$\mathbb{P}(\tau > (1 - 6\epsilon)^{-1}(1 - 2\epsilon)^{-1}(\log n + C) \mid 2w_0 + z_0 \leq \epsilon) \leq e^{-C/4}$$

Proof. Let $\psi_t = (w_t, z_t)^\top$. We may assume $2w + z \leq 2\epsilon$ so that $v = 1 - (w + z) \geq 1 - 2\epsilon$. Define the 2×2 matrices $Q = (-1, 0; 2, -1)$ and $B = (1, 1; 1, 1)$. Computing,

$$\begin{aligned} \mu(w) &= -wv + (w + z)z - n^{-1}z \\ \mu(z) &= (2w - z)v - (w + 2z)z + 2n^{-1}z \end{aligned}$$

So, if $n \geq 1/\epsilon$ then letting $\delta = 2\epsilon/(1 - 2\epsilon)$,

$$\mu(\psi) \leq v(Q + \delta B)\psi$$

Time-change by v^{-1} so that $\mu(\psi) \leq Q_\delta \psi$ with $Q_\delta := Q + \delta B$. Let $s_t = e^{-Q_\delta t} \psi_t$ so that $s_{t \wedge \tau}$ is a non-negative supermartingale. Using non-negativity and optional stopping, we see that

$$e^{-Q_\delta t} \mathbb{E}[\psi_t; \tau > t] = \mathbb{E}[s_t; \tau > t] = \mathbb{E}[s_{t \wedge \tau}; \tau > t] \leq \mathbb{E}[s_\tau] \leq \mathbb{E}[s_0] = \mathbb{E}[\psi_0]$$

and so

$$\mathbb{P}(\tau > t) \leq \mathbb{P}(\tau > t, \ z_t \geq n^{-1}) \leq n \mathbb{E}[z_t; \tau > t] \leq n |e^{Q_\delta t}| \max(w_0, z_0)$$

Now, Q_δ has eigenvalues $-1 + \delta \pm 2\delta^2$ and corresponding eigenvectors $(1, \pm 2\delta)^\top$. Thus $Q_\delta = SAS^{-1}$ with $S = (1, 1; 2\delta, -2\delta)$ and $S^{-1} = (1/2\delta, 1/4; 1/2\delta, -1/4)$. If $\delta \leq 1$, then $|S| \leq 1$, $|S^{-1}| \leq (2\delta)^{-1}$ and $|A| \leq -1 + \delta + 2\delta^2$ and so

$$|e^{Q_\delta t}| \leq |S| |e^{At}| |S^{-1}| \leq (2\delta)^{-1} e^{-(1-\delta-2\delta^2)t}$$

If $\epsilon \leq 1/6$ then $1 - 2\epsilon \geq 2/3$, $\delta \leq 3\epsilon \leq 1/2$ and $1 - \delta(1 + 2\delta) \geq 1 - 6\epsilon$. If $w_0, z_0 \leq \epsilon$ then

$$|e^{Q_\delta t}| \max(w_0, z_0) \leq (\epsilon/2\delta) e^{-(1-6\epsilon)t} = (1 - 2\epsilon) e^{-(1-6\epsilon)t/4}$$

Recalling the time change and noting $v^{-1} \leq (1 - 2\epsilon)^{-1}$, then letting $t = (1 - 6\epsilon)^{-1}(1 - 2\epsilon)^{-1}(\log n + C)$ and using the fact that $1 - 2\epsilon \leq 1$ gives the result. \square

Proposition 11. *Let $\tau = \inf\{t : u_t = 1 \text{ or } u_t \leq 1 - 2\epsilon\}$. For $\alpha > 0$, there is $\epsilon_0 > 0$ so that for $\epsilon \in (0, \epsilon_0]$,*

$$\lim_{n \rightarrow \infty} \sup_{(u, z): u \geq 1 - \epsilon} \mathbb{P}_{(u, z)}(\tau > (1 + \alpha) \log n \quad \text{or} \quad u_\tau \leq 1 - 2\epsilon) = 0$$

and

$$\lim_{n \rightarrow \infty} \inf_{(u, z): |u - (1 - \epsilon + n^{-1})| \leq n^{-1}} \mathbb{P}_{(u, z)}(\tau > (1 - \alpha) \log n) = 1$$

Proof. Since $2w + z = 1 - u$, the above definition of τ agrees with the one used in Lemmas 20-21. Given $\alpha > 0$, $0 < \epsilon_0 < 1/6$ be small enough that for $\epsilon \in (0, \epsilon_0]$,

$$(1 - 6\epsilon)^{-1}(1 - 2\epsilon)^{-1} \leq 1 + \alpha/2 \quad \text{and} \quad (1 + 13\epsilon/2)(1 - \alpha) \leq 1 - \alpha/2.$$

Letting $T = n^{1/2}$ and using the first result of Lemma 20,

$$\mathbb{P}(\tau \leq n^{1/2} \quad \text{and} \quad u_\tau \leq 1 - 2\epsilon \mid u_0 \geq 1 - \epsilon) = o(1) \quad (32)$$

On the other hand, letting $C = (\alpha/2) \log n$ and using the first result of Lemma 21,

$$\mathbb{P}(\tau > (1 + \alpha) \log n \mid u_0 \geq 1 - \epsilon) \leq n^{-\alpha/2}/4 = o(1)$$

Since $n^{1/2} > (1 + \alpha) \log n$ for n large enough, the first statement follows. For the second statement, letting $C = 1 - \alpha$ in Lemma 20,

$$\mathbb{P}(u_\tau = 1 \quad \text{and} \quad \tau \leq (1 - \alpha) \log n \mid u_0 \in [1 - \epsilon, 1 - \epsilon + 2n^{-1}]) \leq 12\epsilon^{-1}n^{-\alpha/2} = o(1)$$

Combining with (32), the second statement follows. \square

Proof of Theorem 2. Let $\tau = \inf\{t : u_t = 1\}$. Recall that $z^* = -2 + \sqrt{5}$. To show the upper bound, for any $\alpha > 0$ take $\epsilon > 0$ small enough to satisfy all conditions, then apply Propositions 9, 10 and 11 in sequence, stopping the process when $u_t \geq 2\epsilon$ and $u_t \geq 1 - \epsilon$, to find that

$$\lim_{n \rightarrow \infty} \sup_{(u,z)} \mathbb{P}_{(u,z)}(\tau > (1 + (2z^*)^{-1} + 3\alpha) \log n) = 0$$

To show the lower bound, in Proposition 9 start from (u, z) achieving the supremum, which is a maximum since the state space is finite. Apply the result of Proposition 9. Then, stop the process when $|u_t - (1 - \epsilon - n^{-1})| \leq n^{-1}$, which occurs before τ since u_t has jumps of size at most $2n^{-1}$. Apply Proposition 11. Combining the two, conclude that for any $\alpha > 0$,

$$\lim_{n \rightarrow \infty} \sup_{(u,z)} \mathbb{P}_{(u,z)}(\tau > (1 + (2z^*)^{-1} - 2\alpha) \log n) = 1$$

\square

Acknowledgements

The author wishes to thank Nicolas Lanchier for suggesting the model, and for many helpful conversations while working on the paper.

Appendix

Miscellaneous estimates.

1. Since

$$\frac{d}{dx}(x^{1+a-\beta}e^{-cx^\beta}) = ((1+a-\beta)x^{-\beta} - c)x^a e^{-cx^\beta}$$

and $t \mapsto c - (1 + a - \beta)t^{-\beta}$ increases with t , if $c - (1 + a - \beta)x^{-\beta} > 0$ we have the upper bound

$$\begin{aligned} \int_x^\infty t^a e^{-ct^\beta} dt &\leq (c - (1 + a - \beta)x^{-\beta})^{-1} \int_x^\infty (c - (1 + a - \beta)t^{-\beta}) t^a e^{-ct^\beta} dt \\ &= (c - (1 + a - \beta)x^{-\beta})^{-1} x^{1+a-\beta} e^{-cx^\beta}. \end{aligned} \quad (33)$$

2. Factoring, using the fact that $|(1 + \lambda)^{-1} \lambda^{1/2+\alpha}/2| \leq 1/2$ and $(1 - x)^{-1} \leq 1 + 2x$ for $|x| \leq 1/2$, then using the fact that $\lambda^{1/2+\alpha} \leq (1 + \lambda)^{1/2+\alpha}$,

$$\begin{aligned} (1 + \lambda - \lambda^{1/2+\alpha}/2)^{-1} &= (1 + \lambda)^{-1} (1 - (1 + \lambda)^{-1} \lambda^{1/2+\alpha}/2)^{-1} \\ &\leq (1 + \lambda)^{-1} (1 + 2(1 + \lambda)^{-1} \lambda^{1/2+\alpha}/2) \\ &\leq (1 + \lambda)^{-1} + (1 + \lambda)^{-3/2+\alpha} \end{aligned} \quad (34)$$

Lemma 22. *Let X be Poisson with mean λ .*

$$\text{For } 0 < x \leq \lambda^{1/2}, \quad \mathbb{P}(X < \lambda - x\lambda^{1/2}) \leq e^{-x^2/2} \quad \text{and} \quad \mathbb{P}(X > \lambda + x\lambda^{1/2}) \leq e^{-x^2/3}. \quad (35)$$

Proof. We have

$$\mathbb{E}[e^{\theta X}] = \sum_{k \geq 0} e^{\theta k} e^{-\lambda} \lambda^k / k! = e^{-\lambda} \sum_{k \geq 0} (\lambda e^{\theta})^k / k! = \exp(\lambda(e^{\theta} - 1)).$$

Also,

$$\mathbb{P}(e^{\theta X} \geq e^{\lambda \theta c}) = \begin{cases} \mathbb{P}(X \geq c\lambda) & \text{if } \theta > 0 \\ \mathbb{P}(X \leq c\lambda) & \text{if } \theta < 0. \end{cases}$$

Using Markov's inequality,

$$\mathbb{P}(e^{\theta X} \geq e^{\lambda \theta c}) \leq e^{-\lambda \theta c} \mathbb{E}[e^{\theta X}] = \exp(\lambda(e^{\theta} - 1 - \theta c))$$

Optimizing in θ gives $\theta = \log c$ which is positive for $c > 1$ and negative for $c < 1$, and

$$\gamma(c) = e^{\theta} - 1 - \theta c = c - 1 - c \log c.$$

Expanding $\gamma(1 + \delta)$ in an alternating Taylor series around $\delta = 0$,

$$\begin{aligned} \gamma(1 + \delta) &\leq -\delta^2/2 + \delta^3/6 \quad \text{for } |\delta| < 1, \text{ so} \\ &\leq \begin{cases} -\delta^2/2 & \text{for } -1 < \delta \leq 0 \\ -\delta^2/3 & \text{for } 0 \leq \delta < 1, \end{cases} \end{aligned}$$

using $\delta^3 \leq \delta^2$ for $\delta \in [0, 1)$ and $\frac{1}{2} + \frac{1}{6} = \frac{2}{3}$. (35) follows for $0 < x < \lambda^{1/2}$ by letting $\delta = x\lambda^{-1/2}$. For $x = \lambda^{1/2}$ it follows by continuity of probability. \square

Lemma 23. *Let (N_t) be a Poisson process with intensity λ . Fix $\alpha \in (0, 1/2]$ and let*

$$\begin{aligned} \tau_1 &= \sup\{t : N_t - \lambda t \geq \lambda^{1/2+\alpha}/2\} \quad \text{and} \\ \tau_2 &= \sup\{t : N_t - \lambda t \leq -\lambda^{1/2+\alpha}/2\} \end{aligned}$$

denote the last passage time of N_t above/below the curve $\lambda t \pm (\lambda t)^{1/2+\alpha}/2$, respectively. If $\lambda \geq 1$ and $t^{2\alpha} \geq 6$ then

$$\begin{aligned} \mathbb{P}(\tau_1 > t) &\leq 6t^{1-2\alpha} e^{-(\lambda t)^{2\alpha}/3} \quad \text{and} \\ \mathbb{P}(\tau_2 > t) &\leq 4t^{1-2\alpha} e^{-(\lambda t)^{2\alpha}/2}. \end{aligned}$$

Proof. Let f denote the function defined by $f(t) = \lambda t + (\lambda t)^{1/2+\alpha}$. Using Lemma 22, for each $t > 0$,

$$\mathbb{P}(N_t > f(t)) \leq e^{-(\lambda t)^{2\alpha}/3}.$$

Since $|f'(t)| \leq 2\lambda$ for any $t \geq 0$, f is Lipschitz with constant 2λ . Using this and the fact that $t \mapsto N_t$ is non-decreasing,

$$\left\{ \sup_{s \in [t-1, t]} N_s - f(s) > 2\lambda \right\} \subseteq \{N_t > f(t)\},$$

so taking a union bound over $t \in \{T+1, T+2, \dots\}$,

$$\mathbb{P}(\sup_{t \geq T} N_t - f(t) > 2\lambda) \leq \sum_{k \geq 1} e^{-(\lambda(T+k))^{2\alpha}/3} \leq \int_T^\infty e^{-(\lambda t)^{2\alpha}/3} dt.$$

Using (33),

$$\int_T^\infty e^{-(\lambda t)^{2\alpha}/3} dt \leq (\lambda^{2\alpha}/3 - (1 - 2\alpha))T^{-(2\alpha)-1} T^{1-2\alpha} e^{-T^{2\alpha}/3}.$$

If $\lambda \geq 1$ and $T^{2\alpha} \geq 6$, this is at most $6T^{1/2-\alpha} e^{-(\lambda T)^{1/2+\alpha}/3}$. An analogous estimate applies for the lower bound, giving 4 instead of 6 and $1/2$ instead of $1/3$ in the exponent, when $\lambda \geq 1$ and $T^{2\alpha} \geq 4$. \square

Lemma 24. *Let X be a non-decreasing quasi-absolutely continuous semimartingale on \mathbb{R}_+ with jump size at most c and defined for $t < \zeta$, where $\zeta = \sup_{r>0} \inf\{t : X_t \geq r\}$ is the first time of explosion. Suppose that*

$$\mu_t(X) \leq b(t) + \ell(t)X_t \quad (36)$$

for some locally integrable non-negative deterministic functions $b(t), \ell(t)$. Let $m(t) = \exp(\int_0^t \ell(s)ds)$ and let $Y_t = X_t/(X_0 m(t)) - \int_0^t b(s)/m(s)ds$ denote the rescaled process. Let $\zeta' = \zeta \wedge \inf\{t : m(t) = \infty\}$ and $\beta = \int_0^\infty b(t)/m(t)^2 dt$, and assume $\beta < \infty$. Then, $\zeta \geq \zeta'$ and for $y \geq 2$,

$$\mathbb{P}(\sup_{t < \zeta'} Y_t \geq y) \leq \mathbb{E}[e^{-(y-2)X_0/4c(1+\beta)}].$$

Proof. First we treat the case $X_0 = 1$, so that $Y_t = X_t/m(t) - \int_0^t b(s)/m(s)ds$. Given $y > 0$ define $\tau(y) = \inf\{t : Y_t \geq y\}$, and note that $\tau(y) < \zeta'$. Since $1/m(t) = e^{-\int_0^t \ell(s)ds}$, $(1/m(t))' = -\ell(t)/m(t)$, so using linearity of the drift and Lemma 4,

$$\mu(Y_t) \leq (b(t) + \ell(t)X_t)/m(t) + X_t(-\ell(t)/m(t)) - b(t)/m(t) = 0,$$

which implies $Y^p \leq 0$. Clearly $\sigma_t^2(Y) = (1/m(t))^2 \sigma_t^2(X)$. Since X is non-decreasing, it has finite variation, so in particular $X^c = 0$, $X^m = X^d$ and $\langle X^m \rangle_t = (\sum_{s \leq t} (\Delta X_s)^2)^p$. In addition, $0 \leq \Delta X_s \leq c$, so $(\Delta X_s)^2 \leq c \Delta X_s$. Using this and $\sum_{t \leq s \leq t+r} \Delta X_s \leq X_{t+r} - X_t$, for any t, r ,

$$\langle X^m \rangle_{t+r} - \langle X^m \rangle_t \leq c \left(\sum_{t \leq s \leq t+r} \Delta X_s \right)^p \leq c(X_{t+r}^p - X_t^p)$$

which implies $\sigma_t^2(X) \leq c\mu_t(X)$. Using $\mu_t(X) \leq b(t) + \ell(t)X_t = b(t) + \ell(t)m(t)Y_t$,

$$\sigma_t^2(Y) \leq (1/m(t))^2 c\mu_t(X) = cb(t)/m(t)^2 + (c/m(t))\ell(t)Y_t.$$

Since $Y_t < y$ for $t < \tau(y)$,

$$\langle Y \rangle_{\tau(y)} \leq c \int_0^{\tau(y)} b(s)/m(s)^2 ds + yc \int_0^{\tau(y)} \ell(s)/m(s) ds = c\beta(\tau(y)) + yc\alpha(\tau(y)),$$

the last equality defining $\alpha(t)$ and $\beta(t)$. Taking the antiderivative,

$$\alpha(t) = \int_0^t e^{-\int_0^s \ell(r)dr} \ell(s) ds = 1 - e^{-\int_0^t \ell(s)ds} = 1 - 1/m(t) \leq 1 \quad \text{for all } t \geq 0.$$

Since $Y_0 = 1$, $Y_{\tau(y)} \geq y$ and $Y^p \leq 0$, it follows that for $\lambda > 0$,

$$Y_{\tau(y)} - Y_0 - Y_{\tau(y)}^p - \lambda \langle Y \rangle_{\tau(y)} \geq y - 1 - \lambda c(\beta + y).$$

Using (4) with $a = y - 1 - \lambda c(\beta + y)$, assuming $\lambda c \leq 1/2$ we find

$$P(\sup_{t < \zeta'} Y_t \geq y) \leq e^{-\lambda a}.$$

Optimizing λa gives $\lambda = (y - 1)/(2c(y + \beta))$ and

$$\lambda a \geq (y - 1)^2/(4cy(1 + \beta/y)) \geq (y - 2)/(4c(1 + \beta)),$$

and if $y \geq 1$ the assumption $c\lambda \leq 1/2$ holds. For general X_0 , first condition on X_0 and apply the above to X_t/X_0 , which has jump size c/X_0 . Then, integrate over X_0 to obtain the result.

To see that $\zeta \geq \zeta'$, note that $\{\zeta \geq \zeta'\} \supset \bigcup_y \{\sup_{t < \zeta'} Y_t < y\}$ and that the above estimate implies the latter event has probability 1. \square

References

- [1] Andrea Baronchelli, Luca DallAsta, Alain Barrat, and Vittorio Loreto. Topology-induced coarsening in language games. *Physical Review E*, 73(1):015102, 2006.
- [2] Andrea Baronchelli, Maddalena Felici, Vittorio Loreto, Emanuele Caglioti, and Luc Steels. Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(06):P06014, 2006.
- [3] Andrea Baronchelli, Vittorio Loreto, and Luc Steels. In-depth analysis of the naming game dynamics: the homogeneous mixing case. *International Journal of Modern Physics C*, 19(05):785–812, 2008.
- [4] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of modern physics*, 81(2):591, 2009.
- [5] Luca DallAsta, Andrea Baronchelli, Alain Barrat, and Vittorio Loreto. Nonequilibrium dynamics of language games on complex networks. *Physical Review E*, 74(3):036105, 2006.
- [6] Jean Jacod and Albert Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, second edition, 2003.